MODELING UNCERTAINTY IN CLIMATE CHANGE: A MULTI-MODEL COMPARISON

By

Kenneth Gillingham, William Nordhaus, David Anthoff, Geoffrey Blanford, Valentina Bosetti, Peter Christensen, Haewon McJeon, John Reilly, and Paul Sztorc

September 2015

COWLES FOUNDATION DISCUSSION PAPER NO. 2022

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Modeling Uncertainty in Climate Change: A Multi-Model Comparison[1]

Kenneth Gillingham, William Nordhaus, David Anthoff, Geoffrey Blanford, Valentina Bosetti, Peter Christensen, Haewon McJeon, John Reilly, Paul Sztorc

September 17, 2015

Abstract

The economics of climate change involves a vast array of uncertainties, complicating both the analysis and development of climate policy. This study presents the results of the first comprehensive study of uncertainty in climate change using multiple integrated assessment models. The study looks at model and parametric uncertainties for population, total factor productivity, and climate sensitivity. It estimates the pdfs of key output variables, including CO_2 concentrations, temperature, damages, and the social cost of carbon (SCC). One key finding is that parametric uncertainty is more important than uncertainty in model structure. Our resulting pdfs also provide insights on tail events.

[1] The authors are grateful to the Department of Energy and the National Science Foundation for primary support of the project. Reilly and McJeon acknowledge support by the U.S. Department of Energy, Office of Science. Reilly also acknowledges the other sponsors the MIT Joint Program on the Science and Policy of Global Change listed at http://globalchange.mit.edu/sponsors/all. The Stanford Energy Modeling Forum has provided support through its Snowmass summer workshops. Kenneth Gillingham currently works as a Senior Economist for the Council of Economic Advisers (CEA). The CEA disclaims responsibility for any of the views expressed herein, and these views do not necessarily represent the views of the CEA or the United States government. None of the authors has a conflict of interest to disclose. Kenneth Gillingham and William Nordhaus are corresponding authors (kenneth.gillingham@yale.edu and william.nordhaus@yale.edu).

I. Introduction

A central issue in the economics of climate change is understanding and dealing with the vast array of uncertainties. These range from those regarding economic and population growth, emissions intensities and new technologies, to the carbon cycle, climate response, and damages, and cascade to the costs and benefits of different policy objectives.

This paper presents the first comprehensive study of uncertainty of major outcomes for climate change using multiple integrated assessment models (IAMs). The six models used in the study are representative of the models used in the IPCC Fifth Assessment Report (IPCC 2014) and in the U.S. government Interagency Working Group Report on the Social Cost of Carbon or SCC (US Interagency Working Group 2013). We focus our efforts in this study on three key uncertain parameters: population growth, total factor productivity growth, and equilibrium climate sensitivity. For the estimated uncertainty in these three parameters, we develop estimates of the uncertainty to 2100 for major variables, such as emissions, concentrations, temperature, per capita consumption, output, damages, and the social cost of carbon.

Our approach is a two-track methodology that permits reliable quantification of uncertainty for models of different size and complexity. The first track involves performing model runs over a set of grid points and fitting a surface response function to the model results; this approach provides a quick and accurate way to emulate running the models. The second track develops probability density functions for the chosen input parameters (i.e., the parameter pdfs) using the best available evidence. We then combine both tracks by performing Monte Carlo simulations using the parameter pdfs and the surface response functions.

This methodology provides a transparent approach to addressing uncertainty across multiple parameters and models and can easily be applied to additional models and uncertain parameters. An important aspect of this methodology, unlike virtually all other model comparison exercises, is its replicability. The approach is easily validated because the data from the calibration exercises are relatively compact and are compiled in a compatible format, the surface responses can be estimated independently, and the Monte Carlo simulations can be easily run in multiple existing software packages.

This paper is structured as follows. The next section discusses the statistical considerations underpinning our study of uncertainty in climate change. Section III presents our methodology for the two-track approach, while the next section discusses selection of calibration runs. Section V gives the derivation of the

probability distributions. Section VI gives the results of the model calculations and the surface response functions, and section VII presents the results of the Monte Carlo estimates of uncertainties. We conclude with a summary of the major findings in section VIII. The Appendices provide further background information.

II. Statistical Considerations

A. Background on Uncertainty in Climate Change

Climate change science and policy have focused largely on projecting the central tendencies of major variables and impacts. While central tendencies are clearly important for a first-level understanding, attention is increasingly on the uncertainties in the projections. Uncertainties take on great significance because of the possibility of non-linearities in responses, particularly the potential for triggering thresholds in earth systems, in ecosystem, or in economic outcomes. To be sure, uncertainties have been explored in major reports, such as the IPCC Scientific Assessment Reports from the first to the fifth. However, these have mainly examined differences among models as a tool for assessing uncertainties about future projections. As we indicate below, our results suggest that parametric uncertainty is quantitatively more important than differences across models for most variables.

In recent reviews of climate change, there is an increasing focus on improving our understanding of the uncertainties. For example, in 2010 the Inter-Academy Review of the IPCC, the primary recommendation for improving the usefulness of the report was about uncertainty:

> The evolving nature of climate science, the long time scales involved, and the difficulties of predicting human impacts on and responses to climate change mean that many of the results presented in IPCC assessment reports have inherently uncertain components. To inform policy decisions properly, it is important for uncertainties to be characterized and communicated clearly and coherently. (InterAcademy Council 2010)

In a recent report, the U.S. Congressional Budget Office also voiced its concerns about uncertainty:

> In assessing the potential risks from climate change and the costs of averting it, however, researchers and policymakers encounter pervasive uncertainty. That uncertainty contributes to great

> differences of opinion as to the appropriate policy response, with some experts seeing little or no threat and others finding cause for immediate, extensive action. Policymakers are thus confronted with a wide range of recommendations about how to address the risks posed by a changing climate—in particular, whether, how, and how much to limit emissions of greenhouse gases. (CBO 2005)

The focus on uncertainty has taken on increased urgency because of the great attention given by scientists to tipping elements in the earth system. An influential study by Lenton et al. (2008) discussed important tipping elements such as the large ice sheets, large-scale ocean circulation, and tropical rain forests. Some climatologists have argued that global warming beyond 2 °C will lead to an irreversible melting of the Greenland ice sheet (Robinson et al. 2012). Once uncertainties are fully included, policies will need to account for the probability that paths may lead across tipping points, with particular concern for ones that have irreversible elements.

A further set of questions involves the potential for fat tails in the distribution of parameters, of outcomes, and of the risk of catastrophic climate change. (A fat- or thick-tailed distribution is one where the probability of extreme events declines slowly, so the tail of the distribution is thick. An important example is the power-law or Pareto distribution, in which the variance of the process is unbounded for certain parameter values.)

The issue arises because of the combination of outcomes that are potentially catastrophic in nature and probability distributions with fat tails. The combination of these two factors may lead to situations in which focusing on central tendencies is completely misleading for policy analysis. In a series of papers, Martin Weitzman (see especially Weitzman 2009) has proposed a dramatically different conclusion from standard analysis in what he has called the Dismal Theorem. In the extreme case, the combination of fat tails, unlimited exposure, and high risk aversion implies that the expected loss from certain risks such as climate change is unbounded and we therefore cannot perform standard optimization calculations or cost-benefit analyses.

There are to date many studies of the implications of uncertainty for climate change and climate-change policy or of uncertainty in one or many parameters using a single model. Some notable examples include Reilly et al. (1987), Peck and Teisberg (1993), Nordhaus and Popp (1997), Pizer (1999), Webster (2002), Baker (2005), Hope (2006), Nordhaus (2008), Webster et al. (2012), Anthoff and Tol (2013), and Lemoine and McJeon (2013).

To date, however, the only published study that aims to quantify uncertainty in climate change for multiple models is the U.S. government Interagency Working Group report on the social cost of carbon, which is published in Greenstone et al. (2013) and more extensively described in IAWG (2010). This study used three models, two of which are included in this study, to estimate the social cost of carbon for U.S. government purposes. However, while it did examine uncertainty, the cross-model comparison focused on a single uncertain parameter (equilibrium climate sensitivity) for its formal uncertainty analysis; all other uncertain parameters in the models were left uncertain with the modelers' pdfs. Even with this single uncertain parameter, the estimated social cost of carbon varies greatly. The 2015 social cost of carbon in the updated IAWG (2013) is $38 per ton of CO2 using the median estimate versus $109 per ton of CO2 using the 95 percentile (both in 2007 dollars and using a 3% discount rate), which would imply very different levels of policy stringency. The IAWG analysis also used combinations of model inputs and outputs that were not always internally consistent. Comparison of the uncertainties in a consistent manner in different models is clearly an important missing area of study.

B. Central approach of this study

This project aims to quantify the uncertainties of key model outcomes induced by uncertainty in important parameters. We hope to learn the degree to which there is precision in the point estimates of major variables that are used in major integrated assessment models. Put differently, the research question we aim to answer from this study is: How do major parameter uncertainties affect the distribution of possible outcomes of major outcomes; and what is the level of uncertainty of major outcome variables?

We call this question one of "classical statistical forecast uncertainty." The study of forecasting uncertainty and error has a long history in statistics and econometrics. See for example Clements and Hendry (1998, 1999) and Ericsson (2001). The standard tools of forecasting uncertainty have virtually never been applied to models in the energy-climate-economy areas because of the complexity of the models and the non-probabilistic nature of both inputs and structural relationships.

Key uncertainties that we will examine include both projections and policy outcomes. For example, what are the uncertainties of emissions, concentrations, temperature increases, and damages in a baseline projection? What is the uncertainty in the social cost of carbon? How do uncertainties across models compare with the uncertainties within models generated by parameter uncertainty?

One of the key contributions of this work is that it has the potential to highlight areas where reducing uncertainty will have a high payoff.

C. Uncertainty in a broader context

There are several uncertainties in climate change that face both natural and social scientists and decision makers. Among the important ones are: (1) parametric uncertainty, such as uncertainty about climate sensitivity or output growth; (2) model or specification uncertainty, such as the specification of the aggregate production function; (3) measurement error, such as the level and trend of global temperatures; (4) algorithmic errors, such as ones that find the incorrect solution to a model; (5) random error in structural equations, such as those due to weather shocks; (6) coding errors in writing the program for the model; and (7) scientific uncertainty or error, such as when a model contains an erroneous theory.

This study focuses primarily on the first of these, parametric uncertainty, and to a limited extent on the second, model uncertainty. We focus on the first because there are major uncertainties about several parameters, because this has been a key area for study in earlier approaches, and because it is a type of uncertainty that lends itself most readily to model comparisons. In addition, since we employ six models, the results provide some information about the role of model uncertainty, although we do not develop a formal approach to model uncertainty. We recognize that parameter and model uncertainties are but two of the important questions that arise, but a rigorous approach to measuring the contribution of these uncertainties will make a major contribution to understanding the overall uncertainty of climate change.

From a theoretical point of view, the measures of uncertainty can be viewed as applying the principles of judgmental or subjective probability, or "degree of belief," to measuring future uncertainties. This approach, which has its roots in the works of Ramsey (1931), de Finetti (1937), and Savage (1954), recognizes that it is not possible to obtain frequentist or actuarial probability distributions for the major parameters in integrated assessment models or in the structures of the models. The theory of subjective probability views the probabilities as akin to the odds that informed scientists would take when wagering on the outcome of an uncertain event. For example, suppose the event was population growth from 2000 to 2050. The subjective probability might be that the interquartile range (25%, 75%) was between 0.5% and 2.0% per year. In making the assessment, the scientist would in effect say that it is a matter of indifference whether to bet that the outcome when known would be inside or outside that range. While it is not contemplated that a bet

would actually occur (although that is not unprecedented), the wager approach helps frame the probability calculation.

III. Methodology

A. Overview of our two track approach

In undertaking an uncertainty analysis, the project contemplated two potential approaches. In one approach, each model would do a Monte Carlo simulation in which it would do many runs where the chosen uncertain parameters are drawn from a joint pdf. While potentially feasible for some models, such an approach is excessively burdensome and likely infeasible at the scale necessary to have reliable estimates.

We therefore developed a second approach which we call the "two-track Monte Carlo." This approach separates the model calibration runs from generation of the parameter pdfs and the Monte Carlo estimates. At the core of the approach are two parallel tracks, which are then combined to produce the final results. The first track uses model runs from six participating economic climate change integrated assessment models to develop surface response functions; these runs provide the relationship between our uncertain input parameters and key output variables. The second track develops probability density functions characterizing the uncertainty for each analyzed uncertain input parameter. We combine the results of the two tracks using a Monte Carlo simulation to characterize statistical uncertainty in the output variables.

B. The approach in equations

It will be helpful to show the structure of the approach analytically. We can represent a model as a mapping from exogenous and policy variables and parameters to endogenous outcomes. The models can be written symbolically as follows:

$$Y^m = H^m(z, \alpha, u) \tag{1}$$

In this schema, Y^m is a vector of model outputs for model m; z is a vector of exogenous and policy variables; α is a vector of model parameters; u is a vector of uncertain parameters to be investigated; and H^m represents the model structure. We emphasize that models have different structures, model parameters, and choice of input variables. However, we can represent the arguments of H without reference to models by assuming some are omitted.

The first step in the project is to select the uncertain parameters for analysis. Once the parameters are selected, each model then does selected calibration runs. The calibration runs take as a central set of parameters the base or reference case for each of the models. It then makes several runs that add or subtract specified increments from each of the base values of the uncertain parameters. This produces a set of input and outputs for each model.

More precisely, here is the procedure for the first track of the approach. Each model has a baseline run with base values for each of the uncertain parameters. Denote the base parameter values as $(u^b_{m,1}, u^b_{m,2}, u^b_{m,3})$. The next step determines a grid of deviation values of the uncertain parameters that each model adds or subtracts from the base values of the uncertain parameters. Denote these deviation values as $\Delta^G = (\Delta_{1,1,1}, \Delta_{1,1,2}, ..., \Delta_{5,5,5})$. The Δ^G vector represents 125 = 5 x 5 x 5 deviations from the modelers' base parameter values. So, for example, the vector $\Delta_{1,1,1}$ would represent one of the 125 grid vectors that takes the first value for each uncertain parameter. Suppose that $\Delta_{1,1,1} = (-0.014, -.02, -2)$. Then that calibration run would calculate the outcomes for $Y^m = H^m(z, \alpha, u^b_{m,1} - .014, u^b_{m,2} - .02, u^b_{m,3} - 2)$, where again $u^b_{m,k}$ is the base value for uncertain parameter *k* for model *m.* Similarly, $\Delta_{3,3,3} = (0, 0, 0)$. For that deviation value, the calibration run would calculate the outcomes for $Y^m = H^m(z, \alpha, u^b_{m,1}, u^b_{m,2}, u^b_{m,3})$, which is the model baseline run.

The third step is to estimate surface response functions (SRFs) for each model and variable outcome. Symbolically, these are the following functions:

(2) $$Y^m = R^m(u_1 - u^b_{m,1}, u_2 - u^b_{m,2}, u_3 - u^b_{m,3}) = R^m(u_{m,1}, u_{m,2}, u_{m,3})$$

The SRFs are fit over the observations of the $u_{m,k}$ from the calibration exercises (125 each for the baseline and for the carbon-tax cases). The SRFs are linear-quadratic-interaction equations as described below.

The second track of the project provides us with probability density functions for each of our uncertain parameters, $f^k(u_k)$. These are developed on the basis of external information as described below.

The final step is to estimate the cumulative distribution of the output variables, $G^m(\bar{Y}^m)$. These are the distributions of the outcome variables $\bar{Y}^m$ for

model *m,* where we note that the distributions will differ by model. The distributions are calculated by Monte Carlo methods, for a sample size of *N*:

$$(3) \qquad G^m(\bar{Y}^m) = \left\langle \sum_{n=1}^{N} \left[1 \text{ if } H^m({}^n u_{m,1}, {}^n u_{m,2}, {}^n u_{m,3}) \le \bar{Y}^m, \text{ otherwise} = 0 \right] \right\rangle / N$$

The notation here is that ${}^n u_{m,k}$ is the nth draw of random variable u_k in the Monte Carlo experiment. This unintuitive equation simply states that the cumulative distribution is equal to the fraction of outcomes in the Monte Carlo simulation where the SRF yields a value of the outcome variable that is less than $\bar{Y}^m$. The distribution of outcomes for each variable and model is conditional on the model structure and on the harmonized uncertainty of the uncertain parameters. For a classic study of Monte Carlo methods, see Hammersley and Handscomb (1964).

C. Integrated Assessment Models

The challenge of analysis and policies for global warming is particularly difficult because it spans many disciplines and parts of society. This many-faceted nature also poses a challenge to natural and social scientists, who must incorporate a wide variety of geophysical, economic, and political disciplines into their diagnoses and prescriptions. The task of integrated assessment models (IAMs) is to pull together the different aspects of a problem so that projections, analyses, and decisions can consider simultaneously all important endogenous variables. IAMs generally do not pretend to have the most detailed and complete representation of each included system. Rather, they aspire to have, at a first level of approximation, models that operate all the modules simultaneously and with reasonable accuracy.

The study design was presented at a meeting where many of the established modelers who build and operate IAMs were present. All were invited to participate. After some preliminary investigations and trial runs, six models were able to incorporate the major uncertain parameters into their models and to provide most of the outputs that were necessary for model comparisons. The following is a brief description of each of the six models. Table A5 in the appendix provides further details on each model.

The DICE (Dynamic Integrated model of Climate and the Economy) was first developed around 1990 and has gone through several extensions and revisions. The latest published version is Nordhaus (2014) with a detailed description in Nordhaus and Sztorc (2014). The DICE model is a globally aggregated model that views the economics of climate change from the perspective of neoclassical economic growth theory. In this approach, economies make investments in capital and in emissions

reductions, reducing consumption today, in order to lower climate damages and increase consumption in the future. The special feature of the model is the inclusion of all major elements in a highly aggregated fashion. The model contains about 25 dynamic equations and identities, including those for global output, CO_2 emissions and concentrations, global mean temperature, and damages. The version for this project runs for 60 five-year periods. It can be run in either an Excel version or in the preferred GAMS version. The version used for this study dates from December 2013 and adds loops to calculate the outcomes for different uncertain parameters. The runs were implemented by William Nordhaus and Paul Sztorc.

The FUND model (Climate Framework for Uncertainty, Negotiation, and Distribution) was developed primarily to assess the impacts of climate policies in an integrated framework. It is a recursive model that takes exogenous scenarios of major economic variables as inputs and then perturbs these with estimates of the cost of climate policy and the impacts of climate change. The model has 16 regions and contains explicit representation of five greenhouse gases. Climate change impacts are monetized and include agriculture, forestry, sea-level rise, health impacts, energy consumption, water resources, unmanaged ecosystems, and storm impacts. Each impact sector has a different functional form and is calculated separately for each of the 16 regions. The model runs from 1950 to 3000 in time steps of 1 year. The source code, data, and a technical description of the model are public (www.fund-model.org), and the model has been used by other modeling teams (e.g., Revesz et al. (2014)). FUND was originally created by Richard Tol (Tol, 1997) and is now jointly developed by David Anthoff and Richard Tol. The runs were implemented by David Anthoff.

The GCAM (Global Change Assessment Model) is a global integrated assessment model of energy, economy, land-use, and climate. GCAM is a long-term global model based on the Edmonds and Reilly model (Edmonds and Reilly 1983a, b, c). GCAM integrates representations of the global economy, energy systems, agriculture and land use, with representations of terrestrial and ocean carbon cycles, and a suite of coupled gas-cycle and climate models. The climate and physical atmosphere in GCAM is based on the Model for the Assessment of Greenhouse-Gas Induced Climate Change (MAGICC) (Meinshausen et al. 2011). The global economy in GCAM is represented in 14 geopolitical regions, explicitly linked through international trade in energy commodities, agricultural and forest products, and other goods such as emissions permits. The scale of economic activity in each region is driven by population size, age, and gender as well as labor productivity. The model is dynamic-recursively solved for a set of market-clearing equilibrium prices in all energy and agricultural good markets every 5 years over 2005-2095. The full

documentation of the model is available at a GCAM wiki (Calvin and et al. 2011). GCAM is open-source, but is primarily developed and maintained by the Joint Global Change Research Institute. The model runs were performed by Haewon McJeon.

The MERGE model (Model for Evaluating Regional and Global Effects of greenhouse gas reduction policies) is an integrated assessment model describing global energy-economy-climate interactions with regional detail. It was introduced by Manne et al. (1999) and has been continually developed since; a recently published description is in Blanford et al. (2014). MERGE is formulated as a multi-region dynamic general equilibrium model with a process model of the energy system and a reduced-form representation of the climate. It is solved in GAMS via sequential joint non-linear optimization with Negishi weights to balance inter-regional trade flows. The economy is represented as a top-down Ramsey model in which electric and non-electric energy inputs are traded off against capital and labor and production is allocated between consumption and investment. The energy system includes explicit technologies for electricity generation and non-electric energy supply, with a resource extraction model for fossil fuels and uranium. The climate model includes a five-box carbon cycle and tracks all major non-CO_2 greenhouse gases and non-CO_2 forcing agents explicitly. Temperature evolves as a two-box lag process, where uncertainty about climate sensitivity is considered jointly with uncertainty about the response time and aerosol forcing. The version used for study includes 10 model regions and runs through 2100, with climate variables projected for an additional century. The runs were implemented by Geoffrey Blanford.

The MIT IGSM (Integrated Global Systems Model) was developed in the early 1990's and has been continually updated. It includes a general circulation model of the atmosphere and its interactions with oceans, atmospheric chemistry, terrestrial vegetation, and the land surface. Its economic component represents the economy and anthropogenic emissions. The full IGSM is described in Sokolov et al. (2009) and Webster et al. (2012). The version of the economic component applied here is described in Chen et al. (2015). The earth system component is a simplified general circulation model resolved in 46 latitude bands and 11 vertical layers in the atmosphere with an 11 layer ocean model. The land system includes 17 vegetation types. The economic component is a multi-sector, multi-region applied general equilibrium model, an empirical implementation consistent with neo-classical economic theory. For the current project, the model operates in a recursive fashion in which the economy drives the earth system model but without feedbacks of climate impacts on the economic system. The economic component is solved for 5 year time steps in GAMS-MPSGE and for this exercise was run through 2100. The

earth system component solves on 10 minute time steps (the vegetation model on monthly time steps). The simulations for this exercise were conducted by Y.-H. Henry Chen, Andrei Sokolov, and John Reilly.

The WITCH (World Induced Technical Change Hybrid) model was developed in 2006 (Bosetti et al. 2006) and has been developed and extended since then. The latest version is fully described in Bosetti et al. (2014). The model divides the world into 13 major regions. The economy of each region is described by a Ramsey-type neoclassical optimal growth model, where forward-looking central planners maximize the present discounted value of utility of each region. These optimizations take account of other regions' intertemporal strategies. The optimal investment strategy includes a detailed appraisal of energy sector investments in power-generation technologies and innovation, and the direct consumption of fuels, as well as abatement of other gases and land-use emissions. Greenhouse-gas emissions and concentrations are then used as inputs in a climate model of reduced complexity (Meinshausen et al. 2011). The version used for this project runs for 30 five-year periods and contains 35 state variables for each of the 13 regions, running on the GAMS platform. The runs were implemented by Valentina Bosetti and Giacomo Marangoni.

IV. Choice of uncertain parameters and grid design

A. Choice of uncertain parameters

One of the key decisions in this study was to select the uncertain parameters. The criteria for selection were (at least after the fact) clear. First, each parameter must be important for influencing uncertainty. Second, parameters should be ones that can be varied in each of the models without excessive burden and without violating the spirit of the model structure. Third, the parameters should be ones that can be represented by a probability distribution, either on the basis of prior research or feasible within the scope of this project.

At an initial meeting, an experiment was undertaken in which each of the models was given six uncertain parameters or shocks to test for feasibility. At the end of this initial test experiment, two of the modeling teams decided not to participate because the initial parameters could not be easily incorporated in the model design or because of time constraints. Three of the parameters fulfilled the above-mentioned criteria, and these were the ones that were incorporated in the final set of experiments.

The final list of uncertain parameters were the following: (1) The rate of growth of productivity, or per capita output; (2) the rate of growth of population;

and (3) the equilibrium climate sensitivity (equilibrium change in global mean surface temperature from a doubling of atmospheric CO_2 concentrations).

Additionally, it was decided to do two alternative policy scenarios. One was a "Base" run in which no climate policies were introduced; and the second, labelled "Carbon Tax" (and sometimes "Ampere") introduced a rapidly rising global carbon tax.[2] A run based on carbon prices was selected (instead of quantitative limits) because many models had undertaken similar runs in other model comparison projects, so they were relatively easy to implement.

Several other parameters were carefully considered but rejected. A pulse of emissions was rejected because it had essentially no impact. A global recession was rejected for the same reason. It was hoped to add uncertainties for technology (such as those concerning the rate of decarbonization, the cost of backstop technologies, or the cost of advanced carbon-free technologies), but it proved impossible to find one that was both sufficiently comprehensive and could be incorporated in all the models. Uncertainty about climate damages was excluded because half the models did not contain damages. A final possibility was to analyze policy runs that had quantitative limits rather than carbon prices. For example, some models had participated in model comparisons in which radiative forcings were limited. This approach was rejected because the carbon tax proved easier to define and implement. Additionally, earlier experiments indicated that quantitative limits were often found infeasible, and this would cloud the interpretation of the results.[3]

[2] The Carbon Tax run was selected from the AMPERE model comparisons to reduce the burden on many of the modelers and so that the results from this study can be compared to those from the AMPERE inter-model comparison study (Kriegler et al. 2015). The specific scenario chosen is known in the AMPERE study as "CarbonTax$12.50-increasing." The full AMPERE scenario database can be found online at https://secure.iiasa.ac.at/web-apps/ene/AMPEREDB.

[3] See particularly the results for Energy Modeling Forum 22 reported in a special issue in *Energy Economics* (e.g., see Clarke and Weyant (2009)). Many models found that tight constraints were infeasible for their base runs. A quantitative limit would almost surely have found that large numbers of the 125 scenarios were infeasible for any tight limit on temperature or radiative forcings.

B. Description of uncertain parameters

We next describe the three uncertain parameters contained in the study. It turned out that harmonizing these across models was more complicated than was originally anticipated, as described below.

(1) *The rate of growth of population.* Uncertainty about the rate of growth of population was straightforward. For global models, there was no ambiguity about the adjustment. The uncertainty was specified as plus or minus a uniform percentage growth rate each year over the period 2010-2100. For regional models, the adjustment was left to the modeler. Most models assumed a uniform change in the growth rate in each region.

(2) *The rate of growth of productivity, or per capita output.* The original design had been to include a variable that represented the uncertainty about overall technological change in the global economy (or averaged across regions). The results of the initial experiment indicated that the specifications of technological change differed greatly across models, and it was infeasible to specify a comparable technological variable that could apply for all models. For example, some models had a single production function, while others had multiple sectors.

Rather than attempt to find a comparable parameter, it was decided to harmonize on the uncertainty of global output per capita growth from 2010 to 2100. Each modeler was asked to introduce a grid of changes in its model-specific technological parameter that would lead to a change in per capita output of plus or minus a given amount (to be described in the next section). The modelers were then instructed to adjust that change so that the range of growth rates in per capita GDP from 2010 to 2100 in the calibration exercise would be equal to the desired range.

(3) *The climate sensitivity.* Modeling uncertainty about climate sensitivity proved to be one of the most difficult issues of harmonization across the different models. While all models have modules to trace through the temperature implications of changing concentrations of GHGs, they differ in detail and specification. The major problem was that adjusting the equilibrium climate sensitivity generally required adjusting other parameters in the model that determine the speed of adjustment to the equilibrium; the adjustment speed is sometimes represented by the transient climate sensitivity. This problem was identified late in the process, after the second-round runs had been completed, and modelers were asked to make the adjustments that they thought appropriate. Some models made adjustments in parameters to reflect differences in large climate

models. Others constrained the parameters so that the model would fit the historical temperature record. The differing approaches led to differing structural responses to the climate sensitivity uncertainty, as will be seen below.

C. Grid design

In the first track, the modeling teams provide a small number of calibration runs that include a full set of outputs for a three-dimensional grid of values of the uncertain parameters. For each of the uncertain parameters, we selected five values centered on the model's baseline values. Therefore, for 3 uncertain parameters, there were 125 runs each for the Base and the Carbon Tax policy scenarios.

On the basis of these calibration runs, the next step involved estimating surface-response functions (SRFs) in which the model outcomes are estimated as functions of the uncertain parameters. The hope was that if the SRFs could approximate the models accurately, then they could be used to simulate the probability distributions of the outcome variables accurately. An initial test suggested that the SRFs were well approximated by quadratic functions. We therefore set the range of the grid so that it would span most of the space that would be covered by the distribution of the uncertain parameters, yet not go so far as to push the models into parts of the parameter space where the results would be unreliable.

As an example, take the grid for population growth. The central case is the model's base case for population growth. Each model then uses four additional assumptions for the grid for population growth: the base case plus and minus 0.5% per year and plus and minus 1.0% per year. These would cover the period 2010 to 2100. For example, assume that the model had a base case with a constant population growth rate of 0.7% per year from 2010 to 2100. Then the five grid points for population growth would be constant growth rates of -0.3%, 0.2%, 0.7%, 1.2%, and 1.7% per year. Population after 2100 would have the same growth rate as in the modeler's base case. These assumptions mean that population in 2100 would be $(0.99)^{90}$, $(0.995)^{90}$, 1, $(1.005)^{90}$, and $(1.01)^{90}$ times the base case population for 2100.

For productivity growth, the grid was similarly constructed, but adjusted so that the growth in per capita output for 2100 added -1%, -0.5%, 0%, 0.5%, and 1% to the growth rate in each year for the period 2010-2100.

For the climate sensitivity, the modelers were to add to the baseline equilibrium climate sensitivity -3°C, -1.5°C, 0 °C, 1.5°C, and 3°C. It turned out that the lower end of this range caused difficulties for some models, and for these the

modelers reported results only for the four higher points in the grid or substituted another low value.

In principle, then, for track I each model reported 5 x 5 x 5 model results for both the Base case and the Carbon Tax policy assumptions.

V. Approach for developing probability density functions

A. General considerations

The three uncertain parameters have been the subject of uncertainty analysis in earlier studies. For each parameter, we reviewed earlier studies to determine whether there was an existing set of methods or distributions that could be drawn upon. The desirable features of the distributions is that they should reflect best practice, that they should be acceptable to the modeling groups, and that they be replicable. It turned out that the three parameters used three different approaches, as will be described below.

B. Population

Population growth has been the subject of projections for many years, and numerous groups have undertaken uncertainty analyses for both countries and at the global level. Our review found only one research group that had made long-term global projections of uncertainty for several years, which was the population group at the International Institute for Applied Systems Analysis (IIASA) in Austria. (For a discussion, see O'Neill et al. (2001)). The IIASA demography group is under the direction of demographer Wolfgang Lutz.

The IIASA stochastic projections were developed over a period of more than a decade and are widely used by demographers. The methodology is summarized as follows: "IIASA's projections…are based explicitly on the results of discussions of a group of experts on fertility, mortality, and migration that is convened for the purpose of producing scenarios for these vital rates" (See http://www.demographic-research.org/volumes/vol4/8/4-8.pdf) The latest projections from 2013 (Lutz et al. 2014) are an update to the previous projections from 2007 and 2001 (Lutz et al. 2008), 2001). The methodology is described as follows:

> The forecasts are carried out for 13 world regions. The forecasts presented here are not alternative scenarios or variants, but the distribution of the results of 2,000 different cohort component projections. For these stochastic simulations the fertility, mortality and migration paths underlying the individual projection

> runs were derived randomly from the described uncertainty distribution for fertility, mortality and migration in the different world regions. (Lutz, Sanderson, and Scherbov 2008)

The background methods are described as follows on page 219 of O'Neill et al. (2001):

> The IIASA methodology is based on asking a group of interacting experts to give a likely range for future vital rates, where "likely" is defined to be a confidence interval of roughly 90% (Lutz 1996, Lutz et al. 1998). Combining subjective probability distributions from a number of experts guards against individual bias, and IIASA demographers argue that a strength of the method is that it may be possible to capture structural change and unexpected events that other approaches might miss. In addition, in areas where data on historical trends are sparse, there may be no better alternative to producing probabilistic projections.

For this study, we are aiming for a parsimonious parameterization of population uncertainty. This is necessary because of the large differences in model structure. We therefore selected the uncertainty about global population growth for the period 2010-2100 as the single parameter of interest. We fitted the growth-rate quantiles from the IIASA projections to several distributions, with normal, log-normal, and gamma being the most satisfactory. The normal distribution performed better than any of the others on five of the six quantitative tests of fit for distributions. Based on these results, we therefore decided to recommend the normal distribution for the pdf of population growth over the period.

In addition, we did several alternative tests to determine whether the projections were consistent with other methodologies. One set of tests examines the projection errors that would have been generated using historical data. A second test looks at the standard deviation of 100-year growth rates of population for the last millennium. A third test examines projections from a report of the National Research Council that estimated the forecast errors for global population over a 50-year horizon (see NRC (2000), Appendix F, p. 344). While these all gave slightly different uncertainty ranges, they were similar to the uncertainties estimated in the IIASA study.

On the basis of this review, we decided to use a normal distribution for the growth rate of population based on the IIASA study that has a standard deviation of the average annual growth rate of 0.22 percentage points per year over the period

2010-2100. More details with a background memorandum on the results are available from the authors.

C. Climate Sensitivity

An important parameter in climate science is the equilibrium or long-run response in the global mean surface temperature to a doubling of atmospheric carbon dioxide. In the climate science community, this is called the equilibrium climate sensitivity. With reference to climate models, this is calculated as the increase in average surface temperature with a doubled CO_2 concentration relative to a path with the pre-industrial CO_2 concentration. This parameter also plays a key role in the geophysical components in the IAMs used in this study. In the remainder of this paper, we will follow the convention in the geosciences and call it the equilibrium climate sensitivity (ECS).

Given the importance of the ECS in climate science, there is an extensive literature estimating probability density functions. These pdfs are generally based on climate models, the instrumental records over the last century or so, paleoclimatic data such as estimated temperature and radiative forcings over ice-age intervals, and the results of volcanic eruptions. Much of the literature estimates a probability density function using a single line of evidence, but a few papers synthesize different studies or different kinds of evidence.

We focus on the studies drawing upon multiple lines of evidence. The IPCC Fifth Assessment report (AR5) reviewed the literature quantifying uncertainty in the ECS and highlighted five recent papers using multiple lines of evidence (IPCC 2014). Each paper used a Bayesian approach to update a prior distribution based on previous evidence (the prior evidence usually drawn from instrumental records or a climate model) to calculate the posterior probability density function. Since each distribution was developed using multiple lines of evidence, and in some cases the same evidence, it would be inconsistent to assume that they were independent and simply to combine them. Further, since we could not reliably estimate the degree of dependence of the different studies, we could not synthesize them by taking into account the dependence. We therefore chose the probability density function from a single study and performed robustness checks to using the results from alternative studies cited in the IPCC AR5.

The chosen study for our primary estimates is Olsen et al. (2012). This study is representative of the literature in using a Bayesian approach, with a prior based on previous studies and a likelihood based on observational or modeled data, such as global average surface temperatures or global total heat content. The prior in Olsen et al. (2012) is primarily based on Knutti and Hegerl (2008). That prior is then

combined with output variables from the University of Victoria ESCM climate model (Weaver et al. 2001) to determine the final or posterior distribution.

Olsen et al. (2012) was chosen for the following reasons. First, it was recommended to us in personal communications with several climate scientists. Second, it was representative of the other four studies we examined and falls into the middle range of the different estimates.[4] Third, sensitivity analyses of the effect on aggregate uncertainty of changing the standard deviation of the Olsen et al. (2012) results found that the sensitivity was small (see the section below on sensitivity analyses). Appendix 1 provides more details on Olsen et al. (2012) and also presents a figure comparing this study to the other studies in the IPCC AR5.

Note that the US government used a version of the Roe and Baker distribution calibrated to three constraints from the IPCC for its uncertainty estimates (IAWG 2010). Specifically, the IAWG Report modified the original Roe and Baker distribution to assume that the median value is 3.0 °C, the probability of being between 2 and 4.5 °C is two-thirds, and there is no mass below zero or above 10 °C. The modified Roe and Baker distribution has a higher mean ECS than any of the models (3.5 °C) and a much higher dispersion (1.6 °C as compared to 0.84 °C from Olsen et al. 2012).

The estimated pdf for Olsen et al. (2012) was derived as follows. We first obtained the pdf from the authors. This pdf was provided as a set of equilibrium temperature values and corresponding probabilities. We then explored families of distributions that best approximated the numerical pdf provided. We found that a log-normal pdf fits the posterior distributions extremely well.

To find the parameters of the fitted log-normal pdf, we minimize the squared difference between the posterior density function from Olsen et al. and the log-normal pdf over the support of the distribution (the L_2 or Euclidian norm). In other words, we minimize the sum of the square of the vertical differences between the posterior pdf and a log-normal pdf over all grid points values in the Olsen et al. (2012) distribution.[5] Figure 1 shows the Olsen et al. (2012) pdf, along with the fitted log-normal density function. The fit is extremely close, with the log-normal distribution always within 0.14% of the Olsen et al. (2012) pdf for any grid point value.

---

[4] In tests, we found that the Olsen et al. (2012) distribution is similar to a simple mixture distribution of all five distributions. We calculate this mixture distribution by taking the average probability over all distributions at each temperature increase.

[5] More precisely we minimize over the range of the Olsen et al. distribution, [1.509, 7.4876] °C, with a grid point spacing of 0.1508 °C.

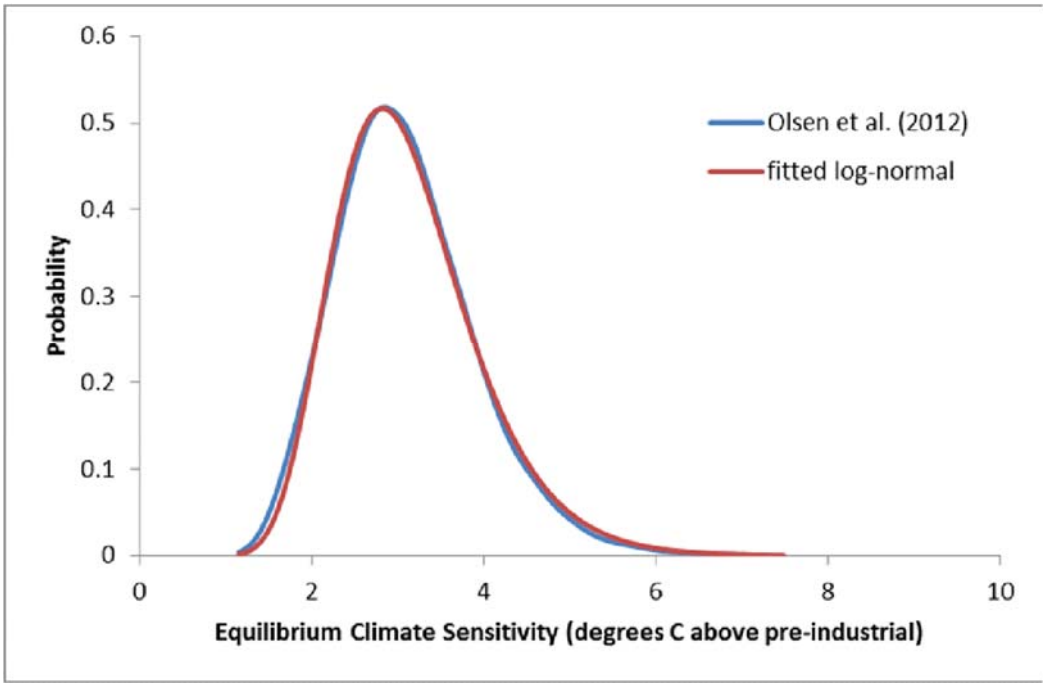


Figure 1. The Olsen et al. (2012) probability density function along with the fitted log-normal distribution used in our analysis.

D. Total Factor Productivity

Uncertainty in the growth of productivity (or output per capita) is known to be a critical parameter in determining all elements of climate change, from emissions to temperature change to damages (Nordhaus 2008). Climate models generally draw their estimates of emissions trajectories from background models of economic growth such as scenarios prepared for the IPCC or studies of the Energy Modeling Forum. No major studies, however, rely on statistically-based estimates of emissions and economic growth.

Forecasts of long-run productivity growth involve active debates on issues such as the role of new technologies and inventions (Brynjolfsson and McAfee 2012, Gordon 2012), potential increases in the research intensity and educational attainment in emerging economies (Fernald and Jones 2014, Freeman 2010), and institutional reform and political stability (Acemoglu et al. 2005). While the empirical literature on economic growth has provided evidence in support of various underlying models, no existing study contains sufficient information to derive a probability distribution for long-run growth rates.

The historical record provides a useful background for estimating future trends. However, it is clear from both theoretical and empirical perspectives that the processes driving productivity growth are non-stationary. For example, estimates of the growth of global output per capita for the 18th, 19th, and 20th century are 0.6, 1.9, and 3.7 percent per year (DeLong 2015 in http://holtz.org/Library/Social%20Science/Economics/Estimating%20World%20GDP%20by%20DeLong/Estimating%20World%20GDP.htm). To the extent that experts on economic growth possess valid insights about the likelihood and possible determinants of long-run growth patterns, then information drawn from experts can add value to forecasts based purely on historical observations or drawn from a single model. Combining expert estimates has been shown to reduce error in short-run forecasts of economic growth (Batchelor and Dua 1995). However, there are few expert studies on long-run growth (see Appendix 2 for discussion) and, to our knowledge, there has been no systematic and detailed published study of uncertainty in long-run future growth rates.

To develop estimates of uncertainties, the project team, led by Peter Christensen, undertook a survey of experts on economic growth to determine both the central tendency and the uncertainty about long-run growth trends. Our survey utilized information drawn from a panel of experts to characterize uncertainty in estimates of global output for the periods 2010-2050 and 2010-2100. We defined growth as the average annual rate of real per capita GDP, measured in purchasing power parity (PPP) terms. We asked experts to provide estimates of the average annual growth rates at 10th, 25th, 50th, 75th, 90th percentiles.

Beginning in the summer of 2014, we sent out surveys to a panel of 25 economic growth experts. As of June 2015, we collected 11 complete results with full uncertainty analysis for the period 2010-2100. A summary of the procedure is provided in Appendix 2, and a complete report will be prepared separately.

There are many different approaches to combining expert forecasts (Armstrong 2001) and aggregating probability distributions (Clemen and Winkler 1999). We assume that experts have information about the likely distribution of long-run growth rates. Their information sets are defined by estimates for 5 different percentiles. We begin by assuming that the estimates are independent across experts and then examined the distributions that best fit the percentiles for each expert and for the combined estimates (average of percentiles) across experts.

We found it useful for this project to characterize the expert pdfs with commonly used distributions so that the Monte Carlo estimates could be easily implemented. In testing the distributions for each expert, we found that most experts' estimates can

be closely fitted by a normal distribution; similarly, the combined distribution is well fitted by a normal distribution. Details are provided in Appendix 2.

The resulting combined normal distribution has a mean growth rate of 2.29% per year and a standard deviation of the growth rate of 1.15% per year over the period 2010-2100. (The mean growth rate of per capita GDP in the base runs of the six models is slightly lower at 1.9% per year over this period.) We test different approaches for combining the expert responses and find little sensitivity to the choice of aggregation method. Figure 2 shows the fitted individual and combined normal pdfs (explained in Appendix 2). In the Monte Carlo estimates below, we chose a standard deviation of the growth rate of per capita output of 1.12% per year (based on the first 11 responses). This value is used in this draft, but will be updated with the addition of further responses.

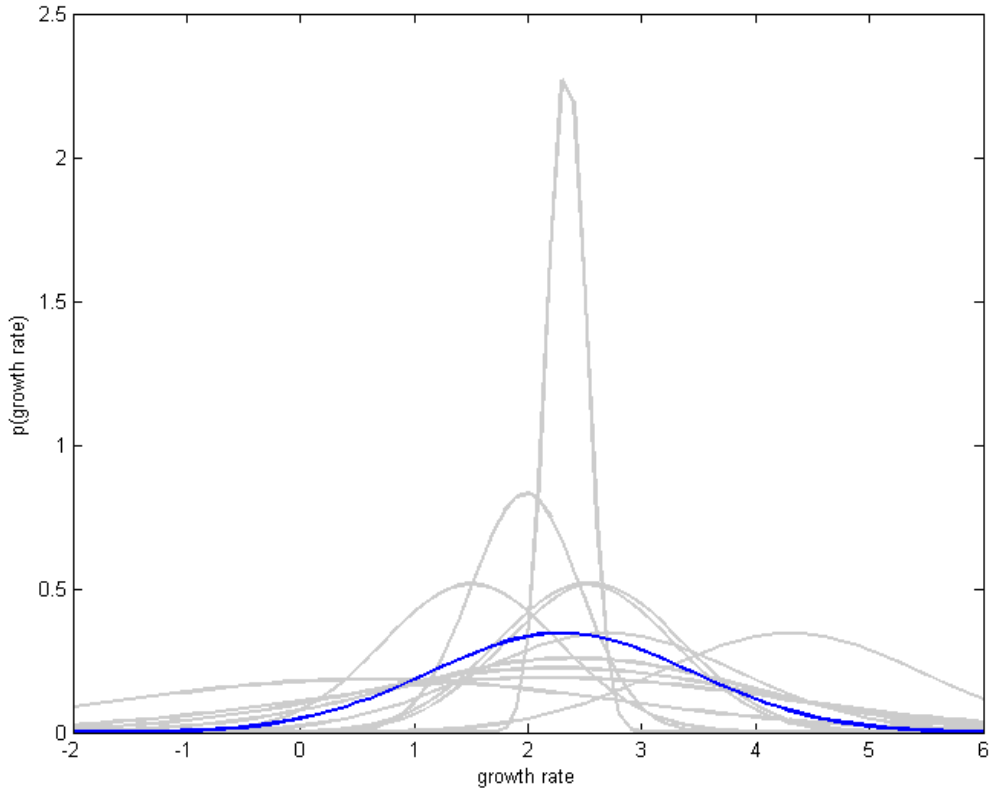


Figure 2. Individual and combined pdfs for annual growth rates of output per capita, 2010 – 2100 (average annual percent per year)

For the methods, see Appendix 2.

It is useful to compare the survey results with historical data. If we take the long-term estimates from Maddison (2003), the 100-year variability of growth over the ten centuries from 1000 to 2000 was 1.5% per year, with a range of -0.1% to 3.7% per year. The variability in these century-step data is higher than the experts' estimate of 1.15% per year.

Global growth rates based on detailed national data are available since 1900. The standard deviation of annual growth rates over this period was 2.9% per year, while the standard deviation of 25-year growth rates was 1.2 or 1.4% per year depending upon the source. The variability of growth in recent years was lower than for the entire period since 1900. The standard deviation in the annual growth rate during the period 1975-2000 was 1.1% per year. We cannot easily translate historical variabilities into century-long variabilities without assuming a specific stochastic structure of growth rates.

VI. Results of Modeling Studies

A. Model results and lattice diagrams

We begin by providing results on the calibration runs and the surface response functions. For each model, there is a voluminous set of inputs and output variables from 2010 to 2100. The full set (consisting of 46,150 x 22 elements) clearly cannot be fully presented. We restrict our focus here to some of the most important results, and consign further results to Appendix 3, with the full results available online at time of publication.

To help visualize the results, we have developed lattice diagrams to show how the results vary across uncertain variables and models. Figure 3 is a lattice diagram for the increase in global mean surface temperature in 2100. Within each of the nine panels, the y-axis is the global mean surface temperature increase in 2100 relative to 1900. The x-axis is the value of the equilibrium temperature sensitivity. Going across panels on the horizontal axis, the first column uses the grid value of the first of the five population scenarios (which is the lowest growth rate); the middle column shows the results for the modeler's baseline population; and the third column shows the results for the population associated with the highest population grid (or highest growth rate).

Going down panels on the vertical axis, the first row uses the highest growth rate for TFP (or the fifth TFP grid point); the middle row shows TFP growth for the modelers' baselines; and the bottom row shows the results for the slowest grid point for the growth rate of TFP. Note that in all cases, the modelers' baseline values

generally differ, but the differences in parameter values across rows or columns are identical.

To understand this lattice graph, begin in the center panel. This panel uses the modeler's baseline population and TFP growth. It indicates how temperature in 2100 across models varies with the ECS, with the differences being 1.5 °C between the ECS grid points. A first observation is that the models all assume that the ECS is close to 3 °C in the baseline. Next, is that the resulting baseline temperature increases for 2100 are closely bunched between 3.75 and 4.25 °C. All curves are upward sloping, indicating a greater 2100 temperature change is associated with a higher ECS.

As the ECS varies from the baseline values, the model differences are distinct. These can be seen in the slopes of the different model curves in the middle panel of Figure 3. We will see below that the impact of a 1 °C change in ECS on 2100 temperature varies by a factor of 2½ across models. For example, DICE, MERGE, and GCAM have relatively responsive climate modules, while IGSM and FUND climate modules are much less responsive to ECS differences. The difference across models becomes larger as we move from the bottom-left to the upper right-hand panel, corresponding to increasing population and TFP growth from bottom left to top right. This result highlights key differences in both the economic and climate components of the different models.

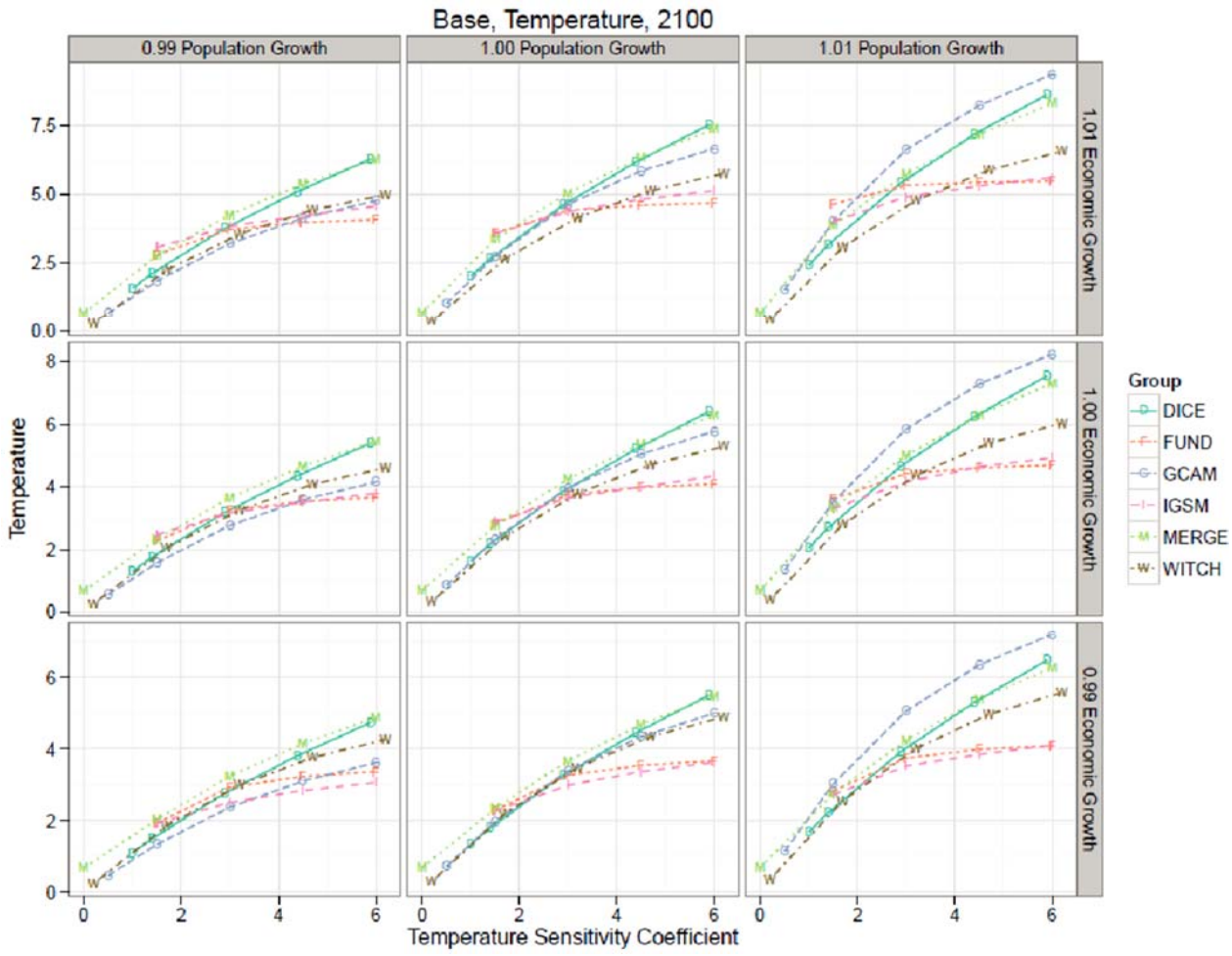


Figure 3. Lattice diagram for 2100 temperature increase
This lattice diagram shows the differences in model results for 2100 global mean surface temperature across population, total factor productivity and temperature sensitivity parameters. The central box uses the modelers' baseline parameters and the Base policy.

Another important relationship to examine is how different models react to the carbon prices. Figure 4 shows the percentage reduction in CO_2 emissions in the Carbon Tax scenario v. the Base run. The horizontal axis shows the magnitude of the carbon tax. One key feature of all models is that attaining zero emissions would require extremely high carbon prices.

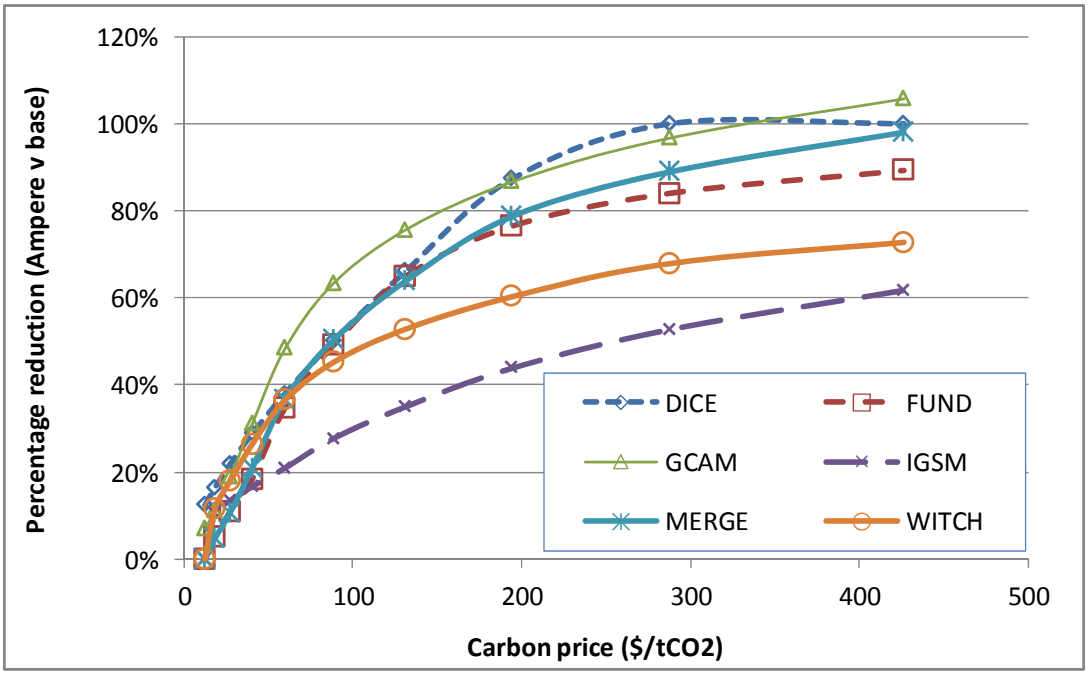


Figure 4. Carbon tax and emissions reductions by model
Models show differing response to higher carbon prices. Note that the carbon prices are all associated with given dates and are common for all models. The points to the far left are for 2010, while the ones at the far right are for 2100. These estimates are for the modelers' baseline parameters.

There are many other results of the modeling exercise. Appendix 3 contains further lattice diagrams, including those for per capita consumption, emissions, and damages, as well as additional tables of results. However, the primary purpose of the present study is to determine the impact of uncertainties, so we leave the model comparisons of major outputs aside at this point.

B. Results of the estimates of the surface response functions

Recall that track I provides the model outcomes (such as output, emissions, and temperature) for each grid-point of a 5 x 5 x 5 x 2 grid of the values of the uncertain parameters and policies. The next step in the analysis is to fit surface response functions (SRFs) to each of the model outputs. These SRFs then will be used, when combined with the Track II probability distributions just discussed, to provide probability distributions of the outcome variables for each model.

We undertook extensive analysis of different approaches to estimating the SRFs. The initial and eventually preferred approach was a linear-quadratic-interactions (LQI) specification. This took the following form:

$$Y = \alpha_0 + \sum_{i=1}^{3} \beta_i u_i + \sum_{j=1}^{3} \sum_{i=1}^{j} \gamma_{ij} u_i u_j$$

In this specification, u_i and u_j are the uncertain parameters. The *Y* are the outcome variables for different models and different years (e.g., temperature for the FUND model for 2100 in the Base run for different values of the 3 uncertain parameters). The parameters α_0, β_i, and γ_{ij} are the estimates from the SRF regression equations. We suppress the subscript for the model, year, policy, and variable.

Table 1 shows a comparison of the results for temperature and log of output for the linear (L) and LQI specifications for the six models. All specifications show marked improvement of the equation fit in the LQI relative to the L version. Looking at the log output specification (the last column in the bottom set of numbers), the residual variance in the LQI specification is essentially zero for all models. For the temperature SRF, more than 99.5% of the variance is explained by the LQI specification. The standard errors of equations for 2100 temperature range from 0.05 to 0.18 °C for different models in the LQI version.

Temperature change, 2100								
	ECS	ECS	TFP	TFP	Pop	Pop	1-R2	1-R2
	L	LQI	L	LQI	L	LQI	L	LQI
DICE	0.9801	1.0403	62.96	60.14	70.40	67.31	0.0239	0.0007
FUND	0.2888	0.4004	58.80	59.83	60.03	61.09	0.0878	0.0053
GCAM	0.9173	0.9744	53.78	52.60	135.05	132.11	0.0788	0.0042
IGSM	0.3296	0.3951	67.28	66.36	48.99	48.18	0.0256	0.0028
MERGE	0.9359	0.9359	57.35	57.35	56.99	56.99	0.0505	0.0039
WITCH	0.8190	0.7883	5.73	28.03	9.69	47.38	0.0911	0.0090

log Output, 2100								
	ECS	ECS	TFP	TFP	Pop	Pop	1-R2	1-R2
	L	LQI	L	LQI	L	LQI	L	LQI
DICE	-0.0411	-0.0371	81.16	81.51	84.95	85.34	0.0018	0.0001
FUND	-0.0016	-0.0023	88.06	88.05	89.72	89.71	0.0000	0.0000
GCAM	0.0000	0.0000	86.74	86.74	88.96	88.96	0.0000	0.0000
IGSM	0.0000	0.0000	89.27	89.27	66.61	66.61	0.0035	0.0002
MERGE	0.0000	0.0000	89.21	89.21	78.44	78.44	0.0006	0.0000
WITCH	0.0000	0.0000	18.38	89.59	17.65	85.63	0.6082	0.0093

The coefficients for ECS are zero in the output equation because there is no feedback from ECS to output in the model.
ECS = temperature sensitivity coefficient
TFP = total factor productivity growth
POP = population growth

Table 1. Linear parameters in of SRF for temperature and log output for linear (L) and liner-quadratic-interactions (LQI) specifications

The linear parameters are the coefficients on the linear term in the SRF regressions. Because the data are decentered (remove the medians), the linear terms in the higher-order polynomials are the derivatives or linear terms at the median values of the uncertain parameters.

The equations are fit as deviations from the central case, so coefficients are linearized at the central point, which is the modelers' baseline set of parameters. Looking at the LQI coefficients for temperature, note that the effect of the ECS on 2100 temperature varies substantially among the models. At the high end, there is close to a unit coefficient, while at the low end the variation is about 0.4 °C per °C in

ECS change. For TFP, the impacts are relatively similar except for the WITCH model, which is much lower. This is likely due to implementation of the TFP changes as input-neutral technical change (rather than changes in labor productivity, as in several other models). For population, the LQI coefficients vary by a factor of three.

For log of output, several models have no feedback from ECS to output and thus show a 0.000 value. The impact of TFP is almost uniform by design. Similarly, the impact of population on output is very similar.

We tested seven different specifications for the SRF: Linear (L), Linear with interactions (LI), Linear quadratic (LQ), Linear, quadratic, linear interactions (LQI) as shown above, 3rd degree polynomial with linear interactions (P3I), 4th degree polynomials with second degree interactions (P4I2), and fourth degree polynomial with fourth degree interactions and polynomial three-way interactions (P4I4S3). For virtually all models and specifications, the accuracy increased sharply as far as the LQI specification. However, as is shown in Figure 5, very little further improvement was found for the more exotic polynomials. In addition to the polynomial interpolations, we investigated several alternative techniques, including Chebyshev polynomials and basis-splines. We found no improvement from these other approaches.

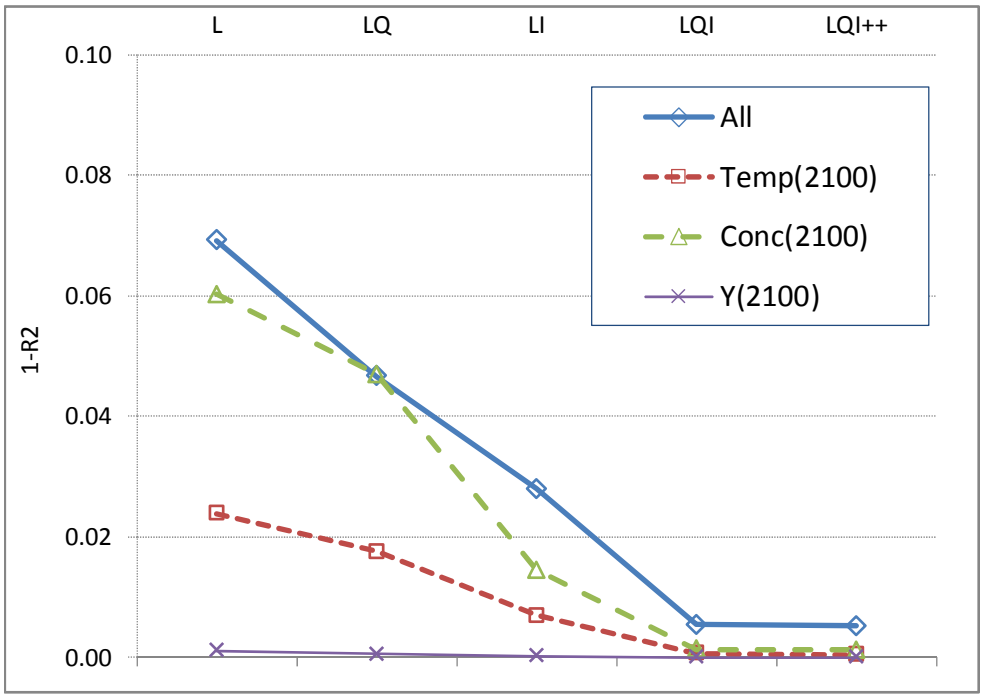


Figure 5. Residual variance for all variables, models, and specifications indicates that for nearly all models, there is little to be gained adding further polynomial terms beyond LQI.

In summary, we found that the linear-quadratic-interaction (LQI) specification of the surface response function performed extremely well in fitting the data in our tests. The reason is that the models, while highly non-linear overall, are generally close to quadratic in the three uncertain parameters. We are therefore confident that they are a reliable basis for the Monte Carlo simulations.

C. Reliability of the MUP procedures with extrapolation

One issue that arises in estimating the distributions of outcome variables is the extent to which the calibration runs in track I adequately cover the range of the pdfs from track II. For both population and the equilibrium temperature sensitivity, the calibration runs cover at least 99.9 % of the range of the pdfs. However, when setting the calibration range for TFP based on earlier informal estimates, we underestimated the variability of the final pdfs. As a result, the calibration runs only extend as far as the 83 percentile at the upper end, requiring us to extrapolate beyond the range of the calibration runs.

Since it was not possible to repeat the calibration runs with an expanded grid, we tested the reliability of the extrapolation and the two track approach with two models. We first examined the reliability for TFP with the base case in the DICE model. This was done by making runs with increments of TFP growth up to 3 estimated standard deviations (i.e., up to a global output growth rate of 6.1% per year to 2100). These runs cover 99.7% of the distribution. We then estimated a surface response function for 2100 temperature over the same interval as for the calibration exercises and extrapolated outside the range. The results showed high reliability of the estimated SRF for temperature increase up to about 2 standard deviations above the baseline TFP growth rate. Beyond that, the SRF tended to overestimate the 2100 temperature. (Similar results were found for CO_2 concentrations and the damage-output ratio in the DICE model.) The reason for the overestimate is that carbon fuels become exhausted at high growth rates, so raising the growth rate further above the already-high rate has a relatively small effects on emissions, concentrations, 2100 temperature, and the damage ratio. Note that this implies that the far upper tail of the temperature distribution using the corrected SRF will show a thinner tail than the one generated by the SRF estimated over the calibration runs.

We also performed a more comprehensive comparison of the MUP procedures with a full Monte Carlo using the FUND model. For this, we took the pdfs for the three uncertain variables and ran a Monte Carlo for the full FUND model with 1 million draws. We then compared the means and standard deviations of different

variables for the two approaches. We tested four different specifications of the SRFs to determine whether these would produce markedly different outcomes. The results indicated that the MUP procedure provided reliable estimates of the means and standard deviations of all variables that we tested except FUND damages. Excepting damages, for the preferred LQI estimate, the absolute average error of the mean for the MUP procedure relative to the FUND Monte Carlo was 0.3%, while the absolute average error for the standard deviation was 1.2%. For damages, the errors were 7% and 44%, respectively. Additionally, the percentile estimates for the MUP procedure (again except for damages) were accurate up to the 90th percentile. And, as will be noted below, the estimates for the parameters of the tails of the distributions were accurate for all variables except damages. A note providing further details on the comparisons is available from the authors.

VII. Results of the Monte Carlo simulations

A. Distributions for major variables

For the Monte Carlo simulations, we took the SRFs for each parameter/model/year/policy and made 1,000,000 draws from each pdf for the three uncertain parameters. We then examined the resulting distributions. This sample size was chosen because the results were reliable at that level. The bootstrap standard errors of the means and the standard deviations were generally less than 0.1% of the mean or standard deviation. The exception was for damages, where the bootstrap standard error of the estimated standard deviations was about 0.2% of the value for the FUND model. We treat each pdf independently, but recognize that there may be some correlation between realizations of population and GDP. However, explorations into this revealed that it did not substantially influence our findings.

Table 2 shows statistics of the distribution of the draws for each of the major outcome variables, with averages taken across all six models. We also show the estimates for the linear and LQI versions to illustrate the sensitivity of the results to the SRF specification. The last column shows the coefficient of variation for each variable. Note that these estimates are within-model (parametric uncertainty) results and do not include across-model variability. The results highlight that emissions, economic output, and damages have the highest coefficient of variation, underscoring that the uncertainty in these output variables is greater than for other variables, such as CO2 concentrations and temperature. This is the result of both the underlying pdfs used and the models themselves.

Variable	Linear				Linear-quadratic-interactions				
	Mean	Standard deviation	10-90 %ile	99 %ile	Mean	Standard deviation	10-90 %ile	99 %ile	Coeff of Variation
CO2 concentrations	888	233	597	1,429	895	247	595	1,672	0.28
Temperature	3.60	0.89	2.26	5.89	3.87	0.89	2.25	6.29	0.23
Output	583	533	1,368	1,825	649	637	1,370	2,975	0.98
Output (log)	664	807	1,343	3,878	664	807	1,343	3,878	1.21
Emissions	112.56	73.10	187.51	282.59	115.12	80.82	187.16	381.98	0.70
Population	12,142	2,378	6,094	17,661	10,245	2,401	6,092	16,816	0.23
Radiative Forcings	7.40	1.60	4.11	11.13	7.40	1.63	4.12	11.81	0.22
Damages	27.41	32.96	84.51	104.36	32.39	41.88	84.90	191.91	1.29
SCC	16.26	7.05	17.68	35.43	13.30	6.95	16.16	36.19	0.52

Table 2. Results of Monte Carlo simulations for averages of all models
The table shows the values of all variables for 2100, except for the social cost of carbon, which is for 2020. Damages and SCC are for three models.

Table 3 shows the percentile distribution for all major variables for all models with results for the base case. The detailed results by models are provided in the appendix. A key result is the distribution of temperature increase for 2100. The median increase across all models is 3.79 °C above 1900 levels. The 95th percentile of the increase is 5.46 °C. Given the size of the interquartile range, these results definitely indicate that there are substantial uncertainties in all aspects of future climate change and its impacts in all the models investigated here.

Variable	0.1 %ile	1 %ile	5 %ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	99.9%ile
CO2 concentrations	533.8	565.1	602.9	633.3	710.5	841.9	1,023.0	1,228.1	1,369.5	1,672.1	2,071.5
Temperature	1.75	2.14	2.55	2.79	3.24	3.79	4.42	5.05	5.46	6.29	7.33
Output	104.3	108.3	114.9	124.7	190.4	422.8	882.5	1,495.0	1,949.1	2,974.5	4,394.6
Output (log)	16.8	38.0	77.6	113.3	211.8	417.6	811.2	1,456.6	2,055.0	3,877.9	7,826.8
Emissions	9.8	18.7	28.5	36.0	55.9	94.2	152.8	223.2	273.1	382.0	528.6
Population	4,961	5,771	6,745	7,352	8,513	10,005	11,717	13,444	14,558	16,816	19,696
Radiative Forcings	3.7	4.3	5.0	5.4	6.2	7.2	8.4	9.6	10.3	11.8	13.7
Damages	(7.3)	(2.8)	(0.1)	1.2	4.3	16.6	44.8	86.1	117.9	191.9	297.2
SCC	2.01	3.49	5.09	6.15	8.40	11.82	16.59	22.31	26.46	36.19	50.82

Table 3. Distribution of all major variables, average of six models
The date for all variables is 2100 except for the SCC, which is 2020. Damages and SCC are for three models.

Table 4 shows the distribution for global temperature increase in 2100 by model. The temperature distributions of the six models are on the whole reasonably close. The median ranges from 3.6 to 4.2°C, with IGSM being the lowest and MERGE being the highest. The interquartile range varies from 0.99 °C (FUND) to 1.39 °C (DICE). The 10-90% ranges from 1.91 °C (WITCH) to 2.65 °C (DICE). Since the variability in the random parameters is the same, the differences are due to model structures.

One interesting feature is the temperature distribution in the tails. The 99th percentile ranges from 5.6 (WITCH) to 7.1 °C (MERGE), while the far tail of the 99.9th percentile ranges from 6.2 (WITCH) to 8.5 °C (MERGE).

Temperature	0.1 %ile	1 %ile	5 %ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	99.9%ile
DICE	1.60	1.97	2.38	2.64	3.12	3.76	4.51	5.29	5.80	6.88	8.28
FUND	1.96	2.30	2.63	2.83	3.19	3.64	4.17	4.74	5.12	5.92	6.96
GCAM	1.59	2.02	2.46	2.73	3.23	3.86	4.56	5.27	5.73	6.64	7.79
IGSM	1.30	1.82	2.31	2.58	3.05	3.58	4.13	4.65	4.97	5.58	6.29
MERGE	2.20	2.56	2.93	3.16	3.61	4.20	4.90	5.63	6.12	7.13	8.46
WITCH	1.83	2.21	2.60	2.82	3.22	3.71	4.23	4.72	5.01	5.58	6.22
Average	1.75	2.14	2.55	2.79	3.24	3.79	4.42	5.05	5.46	6.29	7.33

Table 4. Distribution of temperature change in the Base case, 2100, °C

Table 5 shows the distribution of the SCC for the three models that provide these estimates. These are the estimates of the present value of the flow of future marginal damages of emissions in 2020. Two of the models (WITCH and DICE) use similar quadratic damage functions and are roughly comparable in the middle of the distribution, but the range is much smaller in WITCH.[6] The FUND model has much lower damages (due to a different damage function), and the SCC distribution is an order of magnitude lower than the other two models. Note that the central estimate of the SCC here is $13.30 per ton of CO_2. This is much lower than the preferred estimate of the US government for 2020, which is $46 per ton in 2011$ with a 3% annual discount rate. However, the base case discount rates in the MUP runs for the models that report average 4½% per year to 2050. The IAWG estimate at a 5% discount rate is $13 per ton and therefore consistent with the estimates presented here.

[6] In WITCH multiple regions are modeled, hence the global SCC is the result of the aggregation of regional SCC.

Social cost of carbon	0.1 %ile	1 %ile	5 %ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	99.9%ile
DICE	1.53	3.51	5.81	7.51	11.60	18.36	28.01	39.88	48.72	70.03	103.38
FUND	(0.91)	(0.17)	0.49	0.84	1.38	2.13	3.61	5.56	6.99	10.29	15.03
WITCH	5.41	7.14	8.97	10.09	12.21	14.96	18.15	21.49	23.67	28.24	34.05
Average	2.01	3.49	5.09	6.15	8.40	11.82	16.59	22.31	26.46	36.19	50.82

Table 5. Distribution of social cost of carbon, 2020 (2005$ per ton CO_2)

Figure 6 shows the results for the temperature distributions for the models on a percentile scale. The shapes of the distributions are similar, although they differ by as much as 1 °C in scale across most of the distribution.

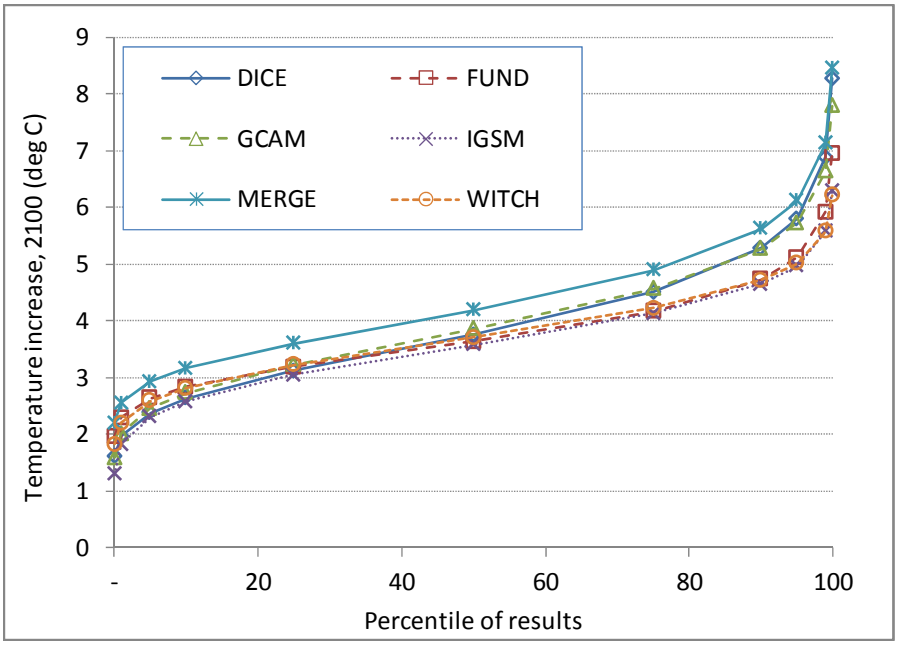


Figure 6. Percentiles of the change in temperature in 2100 across the six models.

An important question that this study can address is whether, based on the current model structures and the assumptions about uncertain parameters, the distributions of outcomes are thin or fat tailed. For these tests, we define a fat tailed distribution as one that has an infinite-variance Pareto or power-law distribution in the tails (based on the discussion in Schuster 1984). Variables with a Pareto distribution have infinite variance when the shape parameter is below 2, and they have an infinite mean with a parameter equal to or less than one. As an informal test, we can examine the ratio of the values of the output variables at the 99th and

99.9th percentile. For a normal distribution, the ratio of these is 1.33. For Pareto distributions with slope values of 2.0, 1.8, and 1.5, the ratios are 3.7, 3.9, and 5.2. If we examine the Monte Carlo estimates, the maximum ratio is 1.56, which occurs for damages in the DICE and FUND models. While this suggests a tail that is slightly fatter than the normal distribution, it falls far short of the slope associated with an infinite-variance Pareto process.

Before presenting the results, we reiterate the concern that the calibration runs do not extend far into the tails for TFP. This implies that the results on tails reported here rely on extrapolations of the SRF outside the sample range. We comment below on our replication of the tail estimates with the FUND model, which are generally accurate. We also emphasize that the estimates of the tails are derived from the interaction of the models with the assumed pdfs. To the extent that the models omit discontinuities or sharp non-linearities, or that our assumed pdfs are too thin-tailed, then we may underestimate the thickness of the tails.

We can also use a formal test of the Pareto shape parameter, although this is more complicated because it requires assumptions about the minimum of the Pareto region (statistical techniques are from Rytgaard 1990). Examining the top 10% of the damage distribution for the DICE model (the most skewed of the variables), we find that the parameter of the Pareto distribution above the 1% right tail is estimated to be 4.7 ($\pm$ 0.047), which is well below the infinite-variance threshold of 2. The Pareto parameter estimate for the 0.1% tail is 7.03 ($\pm$ 0.22). These tests reject the hypothesis that the distributions are fat-tailed in the sense of belonging to an infinite-variance Pareto distribution. The results are due to both the structures of the models and the nature of the shocks. Nothing in the models prevents the generation of fat tails in this situation, but they may miss critical non-linearities, so the tests are not by any means conclusive.

We examined the validity of the results for the tails using the full Monte Carlo estimate of the FUND model discussed above. For these, we compared the informal tests (ratio of the variables at the 99.9%ile to the 99%ile). The MUP calculations were very accurate for all variables except damages, whereas for damages the MUP calculations underestimated the skewness (overestimated the Pareto tail). We also examined the Pareto parameter in the full FUND Monte Carlo and found that the estimate was significantly above the threshold of an infinite variance process.

The results can also be seen in box plots. Figure 7 shows the box plot for temperature increase to 2100. Figure 8 shows the box plot for the CO_2 concentrations for 2100. Both of these underscore that while there are differences between the models in the way that they are run for this study, they are perhaps

smaller than one might have expected – and are much smaller than the within-model variation. We show this formally in the next section.

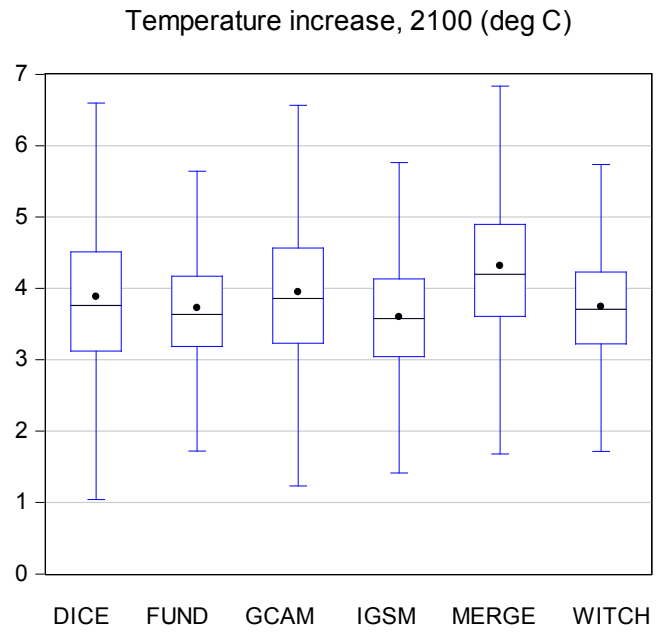


Figure 7. Box plot for the increase in temperature across models in 2100.
Note on boxplots: Dot is mean. Horizontal line is median. Shaded area around line is 95% confidence interval of median (usually too small to see). Box contains interquartile range (IQR or 25 %ile to 75 %ile). The upper staple (horizontal bar) is set at the median plus 2 times the IQR, while lower staple is set at the median minus 2 times the IQR. The upper stable is approximately the 95%ile for most variables. Because of skewness of the variables, the lower staple represents far outliers, and is generally around the 0.1%ile.

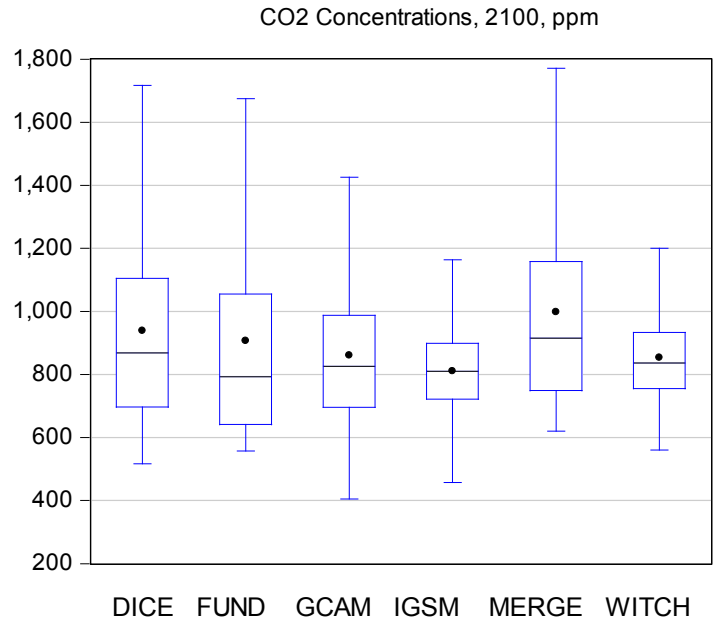


Figure 8. Box plot for CO_2 concentrations, 2100.
For explanation of boxplots, see Figure 7.

B. Model uncertainty v. parametric uncertainty

In examining the uncertainties of climate change and other issues, a common approach has been to look at the differences among forecasts, models, or approaches ("ensembles") and to assume that these are a reasonable proxy for the uncertainties about the end result or endogenous variables. In the area of climate models, for example, researchers have often looked at the equilibrium climate sensitivities in different climate models and assumed that the dispersion would be an accurate measure of the actual uncertainty of the ECS.

It is conceptually clear that the ensemble approach is an inappropriate measure of uncertainty of outcomes. The difference among models represents a measure of structural uncertainty. For example, alternative climate models might have different ways of including cloud feedbacks. Taking all the differences among the models would indicate how state-of-the-art models differ on the processes and variables that they include. Even here, however, existing models are likely to have an incomplete understanding and will therefore underestimate structural uncertainty. However, from a conceptual vantage point, they generally do not

explicitly model and consider parametric uncertainty. In IAMs, to come closer to home, differences in models reflect differences in assumptions about growth rates, production functions, energy systems, and the like. But few models explicitly include parametric uncertainty about these variables. Differences in population growth, for example, are very small relative to measures of uncertainty based on statistical techniques because many models use the same estimates of long-run population trends.

We can use the results of the Monte Carlo simulations to estimate the relative importance of parametric uncertainty and model uncertainty. We can write the results of the Monte Carlo simulations schematically as follows. Assume that the model outcome for variable *i* and model *m* is Y_i^m and that the uncertain parameters are u_i and u_j:

$$Y_i^m = \alpha_i^m + \sum_{i=1}^{3} \beta_i^m u_i + \sum_{j=1}^{3} \sum_{i=1}^{j} \gamma_{i,j}^m u_i u_j$$

For a given distribution of each of the uncertain parameters, the variance of Y_i including model variation is:

$$\sigma^2(Y_i) = \sigma^2(\alpha_i) + \sum_{i=1}^{3} (\beta_i^m)^2 \sigma^2(u_i) + \sum_{j=1}^{3} \sum_{i=1}^{j} (\gamma_{i,j}^m)^2 \sigma^2(u_i)\sigma^2(u_j)$$

The first term on the right hand side is the variance due to model differences (or structural uncertainty), while the second and third terms are the variance due to parameter uncertainty. For this purpose, we include the interaction of the model coefficients (β_i^m and $\gamma_{i,j}^m$) and the parameter uncertainties [$\sigma^2(u_i)$] as parametric uncertainty because they would not be included in ensemble uncertainty. The other terms vanish because we assume that the parametric uncertainties are independent. While dependence will add further terms on the right-hand side of the equation for the variance, it will not affect the fraction due to structural differences due to the first term.

We can easily estimate the total uncertainty and the structural uncertainty for different variables. The results are shown in Table 6. For most variables, virtually all the variance is explained by parametric uncertainty. For example, 94% of the variance of the 2100 temperature increase in all the models is explained by parametric uncertainty, and only 6% is explained by differences in model means. This fact is easily seen in the box charts in Figures 7 and 8. The only variable for

which model uncertainty is important is the social cost of carbon, for which four-fifths of the total variance is due to model differences.

We can put these results in terms of the variabilities due to different factors. If we take the calculated temperature increase to 2100, the overall standard deviation is 0.84 °C including both model and parametric uncertainty. The standard deviation of the model means alone is 0.21 °C. So the variability measured in terms of standard deviations of the temperature increase is underestimated by a factor of four using the ensemble technique.

The net effect of these results is sobering. They indicate that the technique of relying upon ensembles as a technique for determining the uncertainty of future outcomes is (at least for the major climate change variables) highly deficient. Ensemble uncertainty tends to underestimate overall uncertainty by a significant amount.

Variable	Fraction explained by model differences
CO2 concentrations (2100)	0.052
Temperature (2100)	0.061
Output (2100)	0.016
Radiative forcing (2100)	0.020
Population (2100)	0.109
Social cost of carbon (2020)	0.832

Table 6. Fraction of uncertainty (variance) explained by model differences.

C. Sensitivity of the results to parameter variability

An important question is the extent to which the results are sensitive to the individual pdfs for the uncertain parameters. To test for sensitivity, we performed an experiment where we increased the standard deviation of each of the pdfs by a factor of 2, both one at a time and together. For a doubling of the standard deviation of all parameters, the increase in the standard deviation of 2100 temperature was a

factor of 1.83 for all models together. We believe that this is less than two because the short-run temperature impact is not proportional to the ECS.

Table 7 shows the results changing the uncertainty by a factor of two one parameter at a time for the average of the 6 models for all variables which are produced by the six models. The number shows the ratio of the standard deviation of the 2100 value of the variable in the sensitivity case relative to the case with assumed pdfs. Doubling all uncertainties produces close to a doubling of the output uncertainty, with some deviations because of non-linearities.

Doubling population uncertainty has a small effect on all variables except population. Doubling equilibrium temperature uncertainty raises the uncertainty of 2100 temperature by 40% but has no significant effect on other uncertainties. The major sensitivity is TFP uncertainty. Doubling this uncertainty leads to close to doubling of the uncertainty of other major economic variables, and to an increase of 62 percent in the uncertainty of 2100 temperature. This result is similar to a result in van Vuuren et al. (2008), which suggests that uncertainty in GDP growth dominates the uncertainty in emissions.

Variation	CO2 Conc	Temp	Output	Emissions	Population	Rad Forc
Base	1.00	1.00	1.00	1.00	1.00	1.00
Pop x 2	1.11	1.06	1.07	1.11	2.07	1.12
TFP x 2	2.16	1.62	2.68	2.23	1.00	1.99
ETS x 2	1.00	1.40	1.00	1.00	1.00	1.01
All x 2	2.24	1.97	2.74	2.31	2.06	2.07

Table 7. Sensitivity of outcomes for changes in standard deviation of each uncertain parameter by factor of 2

The figure gives the ratio of the standard deviation of the variable at the top of the column to the standard deviation in the base run. For example, doubling the standard deviation of population increased the standard deviation of 2100 temperature by 6%.

The summary on sensitivity of the results to the pdfs shows an important and surprising result. On the whole, the results are insensitive to changes in the population growth pdf; are moderately sensitive to the uncertainty about

equilibrium temperature sensitivity on temperature (as well as to damages and the social cost of carbon, not shown); and are extremely sensitive to the uncertainty about the rate of growth of productivity. While long-run productivity growth has the greatest impact on uncertainty, it is also the least carefully studied of any of the parameters we have examined. This result suggests that much greater attention should be given to developing reliable estimates of the trend and uncertainties about long-run productivity.

VIII. Conclusions

This study is the first multi-model analysis of parametric uncertainty in economic climate-change modeling. The approach is based on estimating classic statistical forecast uncertainty. The central methodology consists of two tracks. Track I involves doing a set of model calibration runs for the six models and three uncertain parameters and estimating a surface response function for the results of those runs. Track II involves developing pdfs for key uncertain parameters. The two tracks are brought together through a set of Monte Carlo simulations to estimate the output distributions of multiple output variables that are important for climate change and climate-change policy. This approach is replicable and transparent, and overcomes several obstacles for examining uncertainty in climate change.

Here are the key results. First, the central projections of the integrated assessment models (IAMs) are remarkably similar at the modeler's baseline parameters. This result is probably due to the fact that models have been used in model comparisons and may have been revised to yield similar baseline results. However, the projections diverge sharply when alternative assumptions about the key uncertain parameters are used, especially at high levels of population growth, productivity growth, and equilibrium climate sensitivity.

Second, despite these differences across models for alternative parameters, the distributions of the key output variables are remarkably similar across models with different structures and levels of complexity. To take year 2100 temperature as an example, the quantiles of the distributions of the models differ by less than ½ °C for the entire distribution up to the 95th percentile.

Third, we find that the climate-related variables are characterized by low uncertainty relative to those relating to most economic variables. For this comparison, we look at the coefficient of variation (CV) of the Monte Carlo simulations. As shown in Table 2, CO_2 concentrations, radiative forcings, and temperature (all for 2100) have relatively low CV. Output and damages have relatively high CV. As examples, the model-average coefficient of variation for carbon dioxide concentrations in 2100 is 0.28, while the coefficient of variation for

climate-change damages is 1.29. The social cost of carbon has an intermediate CV within models, but when model variation is included the CV is close to that of output and damages. These results highlight the importance of further research on economic variables and damage functions for reducing uncertainty and improving policymaking (e.g., see Pizer et al. 2014 and Drouet et al. 2015).

Fourth, we find much greater parametric uncertainty than structural (across model) uncertainty for all output variables except the social cost of carbon. For example, in examining the uncertainty in 2100 temperature increase, the difference of model means (or the ensemble uncertainty) is approximately one-quarter of the total uncertainty, with the rest driven by parametric uncertainty. While looking across six models by no means spans the space of methods, the six models examined here are representative of the differences in size, structure, and complexity of IAMs. This result is important because of the widespread use of ensemble uncertainty as a proxy for overall uncertainty and highlights the need for a re-orientation of research towards examining parametric uncertainty across models.

A fifth interesting finding of this analysis is the lack of evidence in support of fat tails in the distributions of emissions, global mean surface temperature, or damages. Population growth, total factor productivity growth, and climate sensitivity are very likely to be three of the key uncertain parameters in climate change. Yet, based on both informal and formal tests, the models as currently constructed find that the tails are relatively thin. The decline in probabilities associated with a change in any of the variables is much larger than would be associated with an infinite-variance Pareto process. As discussed above, we emphasize that these findings should be interpreted in the context of the current group of models and the assumed pdfs. The results do not rule out fat tails, but they do provide empirical evidence against fat tails in outcomes investigated in this study for the current set of models and the distributions of the three uncertain variables considered here. These results tend to support the use of expected benefit-cost analysis for climate change policy, in contrast to suggestions by some authors that neglect of fat tail events may vitiate standard analyses (Weitzman 2009).

Sixth, we find that within a wide range of uncertainty, changes in dispersion of two of the uncertain parameters taken singly have a relatively small effect on the uncertainty of the output variables, these being population growth and equilibrium temperature sensitivity. However, uncertainty about productivity growth has a major impact on the uncertainty of all the major output variables. The reason for this is that the uncertainty of productivity growth from the expert survey compounds greatly over the 21st century and induces an extremely large uncertainty

about output, emissions, concentrations, temperature change, and damages by the end of the century.

As in any study, this analysis is intentionally sharply focused. By analyzing parametric uncertainty in three key parameters, we do not claim to be capturing all uncertainties in climate change. As we describe above, there are many uncertainties that cannot be captured using the statistical framework developed here. But by providing detailed estimates of uncertainty across a range of IAMs that are currently being used in the policy process, we believe that we have significantly improved the understanding of uncertainty in climate change. Moreover, our new two-track methodology is well-suited for expansion to additional parameters and models, and can be readily used to explore additional concerns, such as the interaction between carbon policies and uncertainty.

Appendix 1. Further Details on the Choice of ECS Distribution

This appendix explains the procedure for developing the pdf for climate sensitivity. The study began by reviewing the five probability density functions for equilibrium climate sensitivity (ECS) used in the IPCC AR5 that draw upon multiple lines of evidence. These are Aldrin et al. (2012), Libardoni and Forest (2013), Olsen et al. (2012), Annan and Hargreaves (2006), and Hegerl et al. (2006). Figure A1 illustrates the log-normal fits to each of these distributions (fits by the present authors).

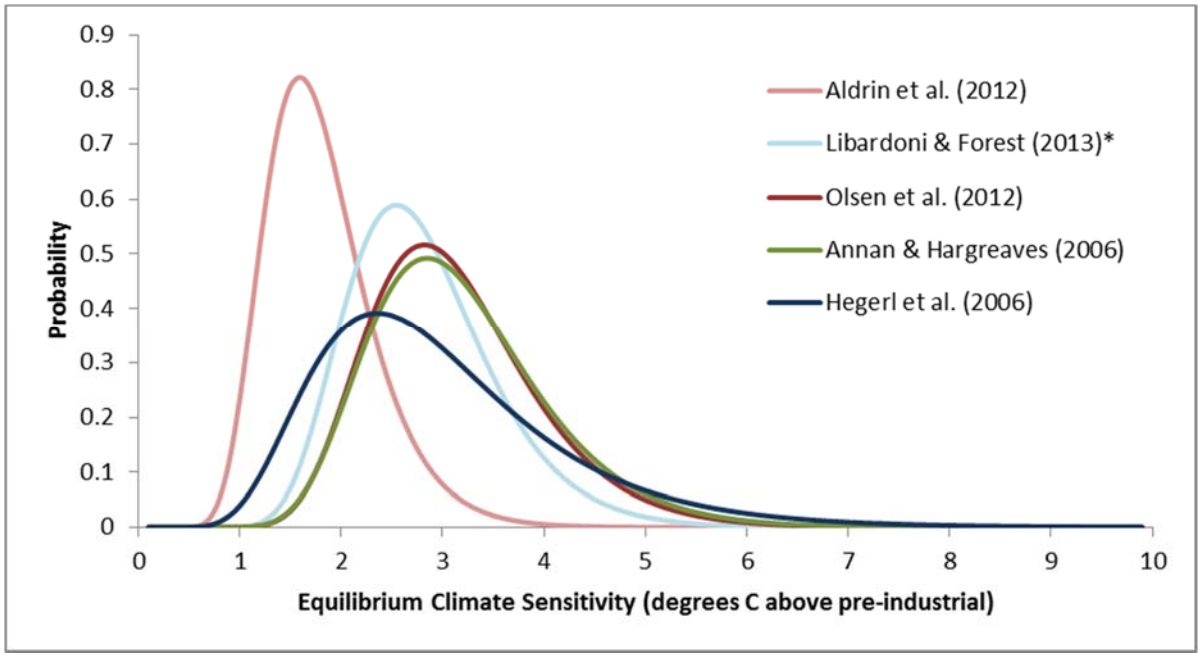


Figure A1. Log-normal distributions fit to the probability density functions cited in the IPCC AR5. The distribution shown here is from the updated Libardoni & Forest (2013) figures.

Our chosen study, Olsen et al. (2012), is representative of the studies in both its methodology and results. It uses a Bayesian approach. The prior distribution was constructed to fit the "most likely" values and "likely" ranges in Figure 3 in Knutti and Hegerl (2008) based on the summary statistics of the "current mean climate state" and "Last Glacial Maximum models." Olsen et al. assume an inverse Gaussian (Wald) distribution and obtain this prior by assuming independence between the

current mean climate state and the last glacial maximum models, and then computing the mixture distribution.

The posterior distribution is then calculated by using a Markov Chain Monte Carlo simulation to update the prior with a likelihood function. The likelihood is based on several different tracers, such as global average atmospheric surface/ocean surface temperatures and global total heat content. These tracers come from the University of Victoria ESCM climate model, which consists of a three-dimensional ocean general circulation model coupled with a thermodynamic/dynamic sea-ice model. The authors assume independence, so that the likelihood of both observations is equal to the product of the likelihoods.

The parameters of the log-normal distribution fit to Olsen et al. are μ = 1.10704 and σ = 0.264. The major summary statistics of the reference distribution in the study are the following: mean = 3.13, median = 3.03, standard deviation = 0.843, skewness = 0.824, and kurtosis = 4.23. In implementing the Monte Carlo for each model, we retained the mean ECS for that model. We then imposed a log-normal distribution that retained the arithmetic standard deviation of the ECS (i.e., a standard deviation of 0.843) based on the Olsen et al. (2012) distribution.

Appendix 2. Expert Survey on Total Factor Productivity

A key component of the project was determining the uncertainty in productivity (or, as operationally defined, output per capita). A review of existing studies indicated that there were no detailed studies of future output uncertainties out to 2100 that we could rely on. We therefore decided to undertake an expert elicitation. The detailed results of the survey will be shortly available separately as a working paper. This appendix sketches the methods and summarizes the preliminary results. Note that the current results include only 11 of the respondents, and the complete survey results will be used for the final publication.

2.1 Survey Design

In determining the probability distribution of future productivity growth, a major difficulty is the non-stationarity of this variable. It is clearly non-stationary if one examines historical data. From a theoretical point of view, we would expect non-stationarity because the major determinants of long-run growth – invention and technological change – involve new and different processes rather than replication of some underlying process. For this reason, it is important to overlay any empirical study with expert views.

Expert opinion has been used systematically in a very limited number of studies of economic growth. For example, Webster et al. (2002) analyze uncertainty in the GDP growth rate out to 2100 (as a proxy for changes in labor productivity) using estimates collected from an elicitation of 5 experts from a single institution. This seemed too thin a base for the present study.

In this study, we conducted a survey of expert predictions about uncertainty in global annual growth rates for the period 2010-2100. Experts provided responses using an online survey (see Figure A2 for the response format). The panel of experts was selected through a process of nomination by leading economists.

We asked experts about growth rates in high-, medium-, and low-income countries, as well as about global aggregate rates. As part of the survey, we alerted experts to problems of overconfidence and include a warm-up section that was designed to increase awareness of their personal overconfidence. In addition, we asked experts about any ambiguities that they experienced in the survey and

provided them with historical data on growth rates for the period 1900-2000 from Barro-Ursua (2010) and Maddison (2003). [7]

Q2. Please provide estimates for the 10th, 25th, 75th, 90th percentile rates of growth of per capita real GDP (average percent per year) for the following regions. The interpretation is that, in your judgment, there is a 1 in 4 chance that the actual value will be less than the 25th percentile and a 9 in 10 chance that the actual value will be less than the 90th percentile. (As noted above, you should feel free not to answer questions for which you think you have no useful information to add)

	2010-2050					2010-2100				
	10th	25th	50th	75th	90th	10th	25th	50th	75th	90th
World										
United States										
China										
High Income (Includes US)										
Middle Income										
Low Income (includes China and India)										

Figure A2. Response Format for Expert Survey

The survey was comprised of 4 sets of questions about growth rates: (1) central estimates (50th percentile) for growth rates for 2010-2050 and 2010-2100, (2) estimates of uncertainty based on providing the 10th, 25th, 75th, and 90th percentiles of the growth rates, (3) the projected magnitude of effects of positive and negative shocks to the economy, and (4) near-term predictions (for the following year). We asked each expert to describe the rationale for their response as well as an explanation of major positive and negative shocks. The survey also asked experts to identify outside sources of information that were used to generate forecasts and to rank their own expertise overall and for particular regions.

2.2 Combining Expert Distributions

We use two methods to estimate the mean and standard deviation for the best-fitting combined normal distribution of growth rates for the period 2010-2100.

The first method assumes that experts have estimates of quantiles of the distribution of long-run growth rates. The combined pdf is then the distribution that minimizes the sum of squared differences between the combined normal

[7] Barro-Ursua Macroeconomic Data available at: rbarro.com/data-sets/. Maddison is from Angus Maddison (2003). Available at: http://www.theworldeconomy.org/statistics.htm.

distribution at each quantile and the average of the quantile estimates of the experts. The second method begins with estimates of the parameters of the best-fitting normal distribution for each expert; and then takes the sample means of the parameters of the experts for the combined normal distribution.

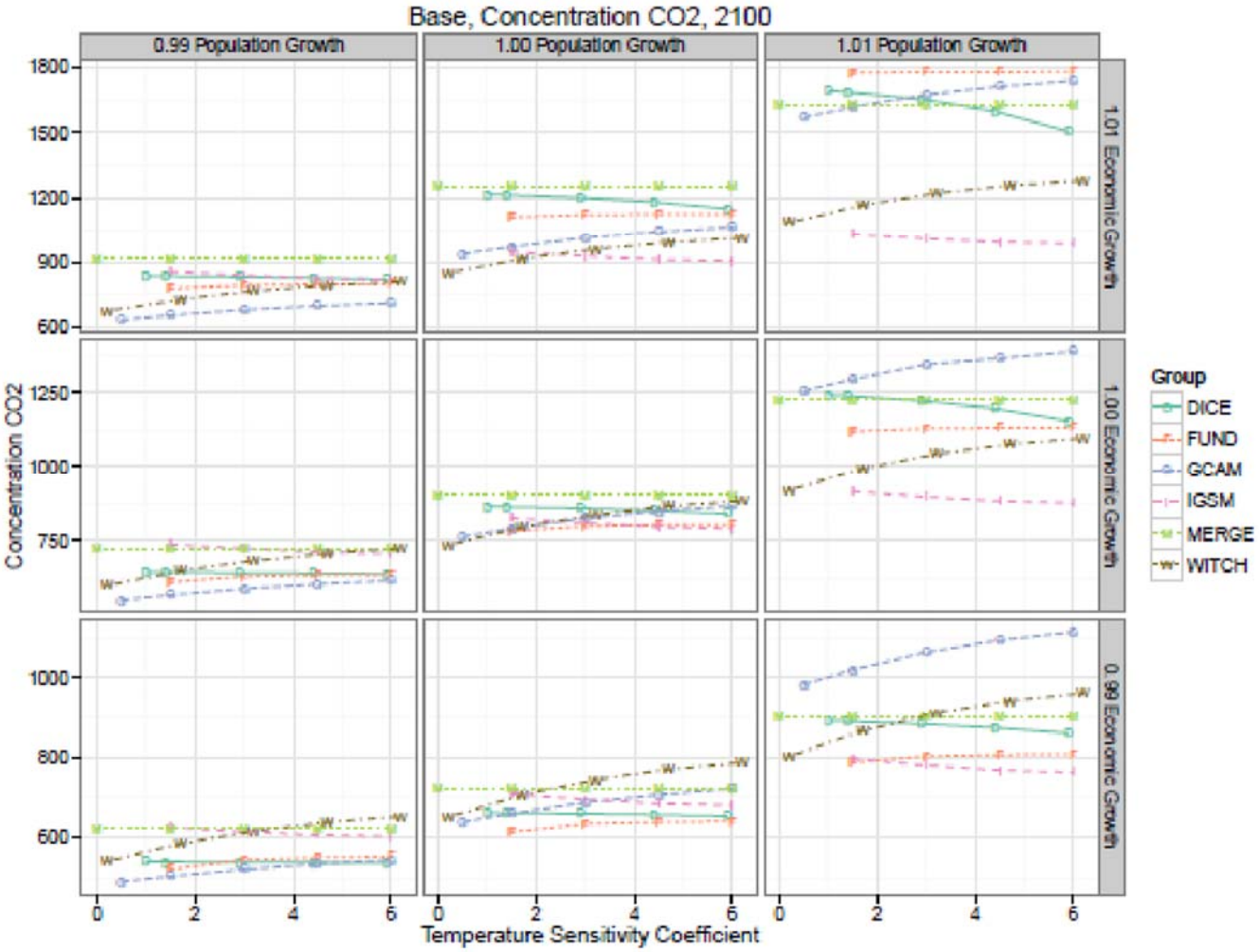
We find very little difference between the two methods. For the preliminary sample, the mean growth rates of per capita output for the two methods are 2.29 and 2.30, respectively for methods 1 and 2. The combined standard deviations are 1.15 and 1.17, respectively.

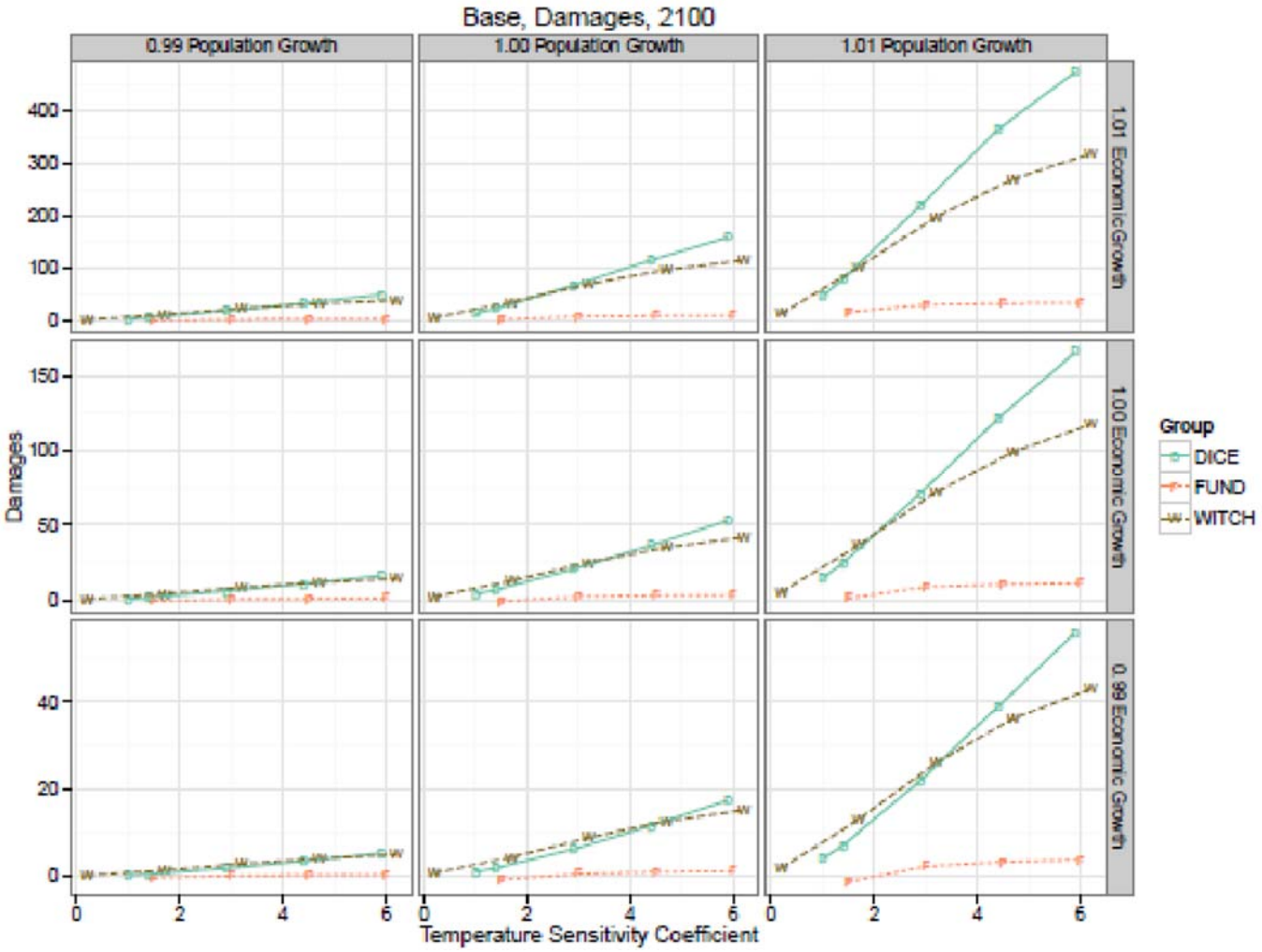
The combined pdfs along with 11 preliminary responses are shown in Figure 2 in the main text. The current procedure uses the sample mean of the standard deviation for the Monte Carlo estimates, but we are considering using a robust estimator for the final report.

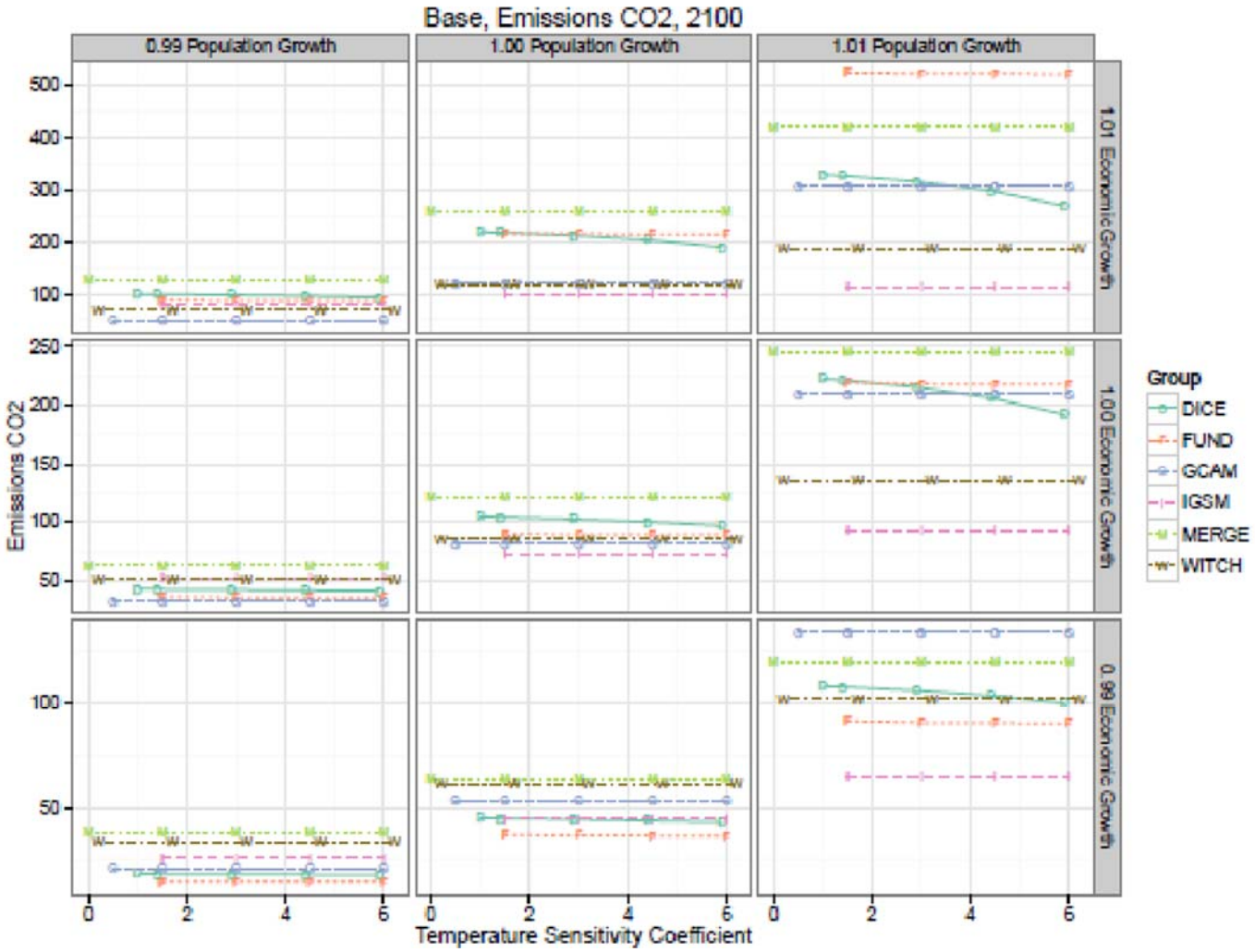
Appendix 3. Additional Lattice Diagrams

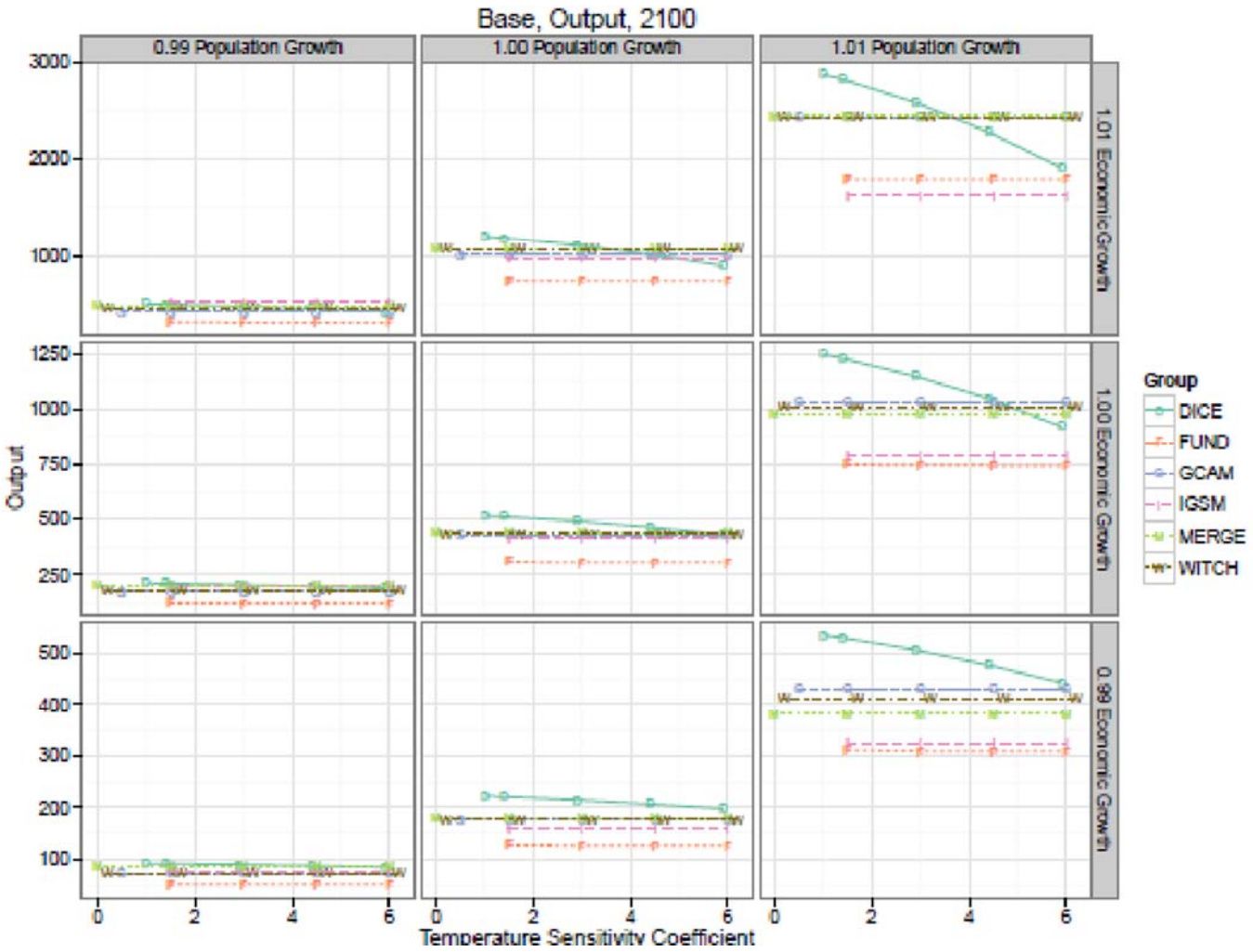
We include here further lattice diagrams. The structure is as described in the text. The only difference is the output variable, which is shown at the top of the graph.

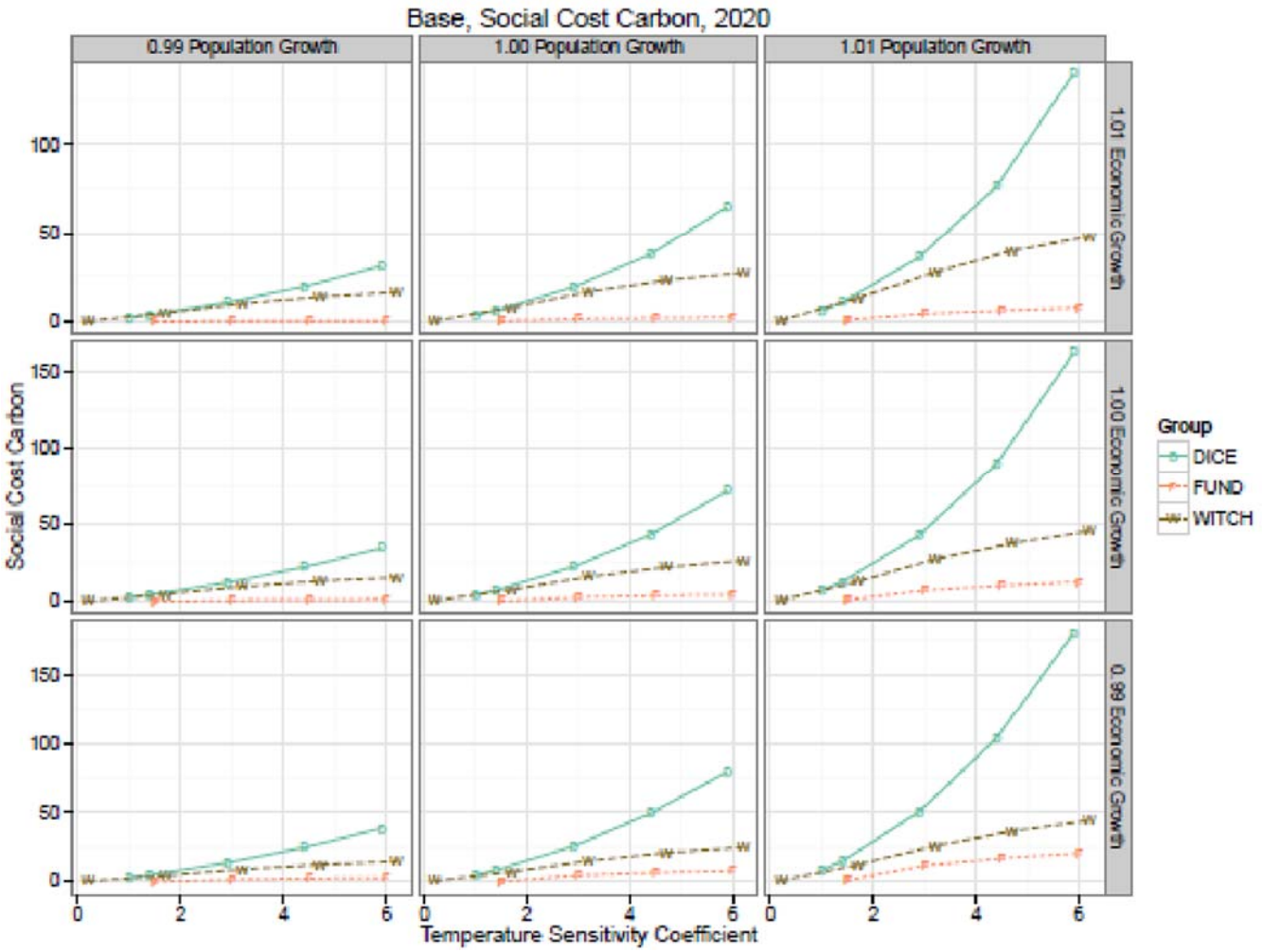
Note that the first group of diagrams is for the base runs, while the second group is for the runs with carbon taxes (Carbon Tax or Ampere runs).

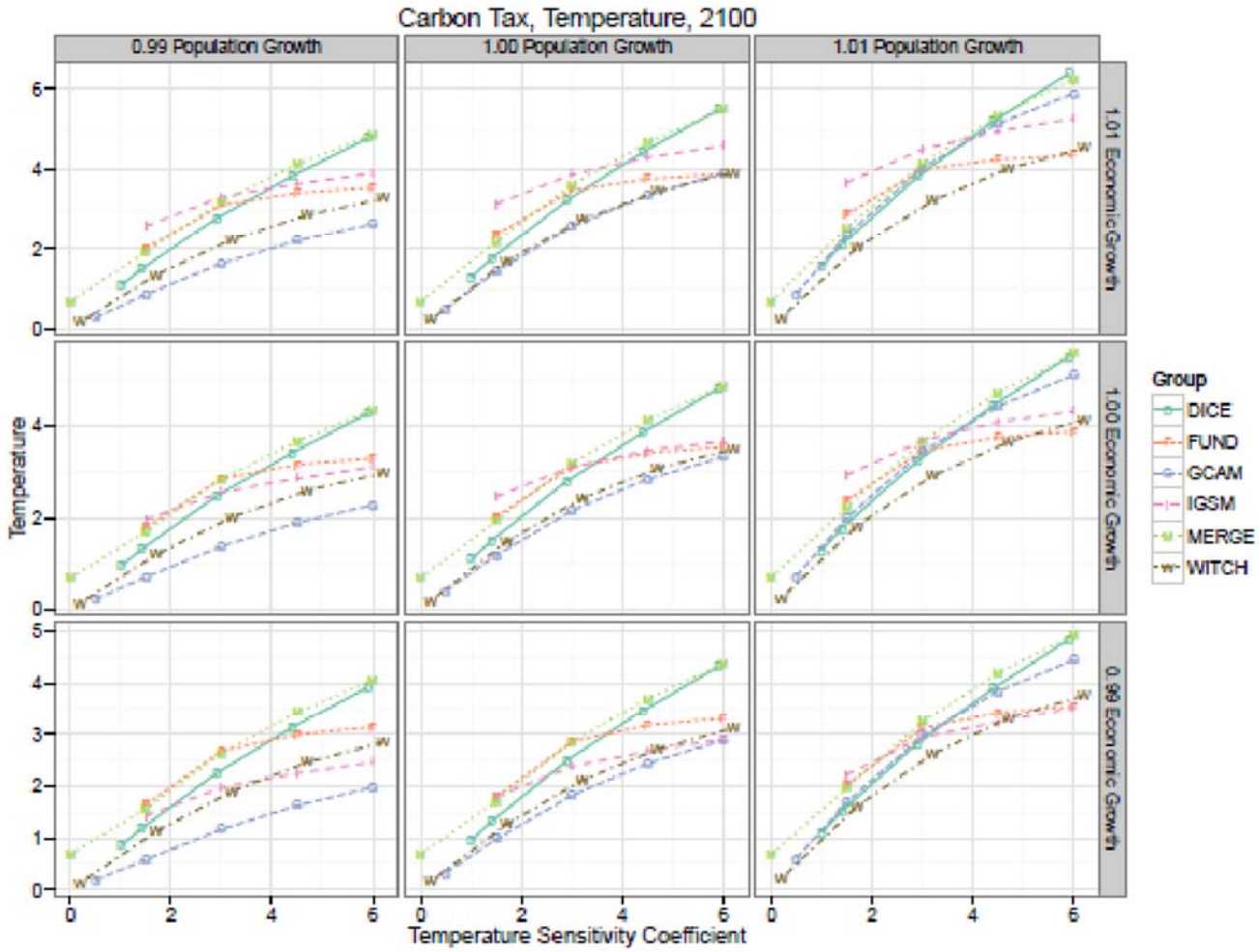



Base, Damages, 2100
0.99 Population Growth
1.00 Population Growth
1.01 Population Growth
1.01 Economic Growth
1.00 Economic Growth
0.99 Economic Growth
Damages
Temperature Sensitivity Coefficient
Group
DICE
FUND
WITCH


Base, Emissions CO2, 2100
0.99 Population Growth
1.00 Population Growth
1.01 Population Growth
1.01 Economic Growth
1.00 Economic Growth
0.99 Economic Growth
Emissions CO2
Temperature Sensitivity Coefficient
Group
DICE
FUND
GCAM
IGSM
MERGE
WITCH


Base, Output, 2100
0.99 Population Growth
1.00 Population Growth
1.01 Population Growth
1.01 Economic Growth
1.00 Economic Growth
0.99 Economic Growth
Output
Temperature Sensitivity Coefficient
Group
DICE
FUND
GCAM
IGSM
MERGE
WITCH


Base, Social Cost Carbon, 2020
0.99 Population Growth
1.00 Population Growth
1.01 Population Growth
1.01 Economic Growth
1.00 Economic Growth
0.99 Economic Growth
Social Cost Carbon
Temperature Sensitivity Coefficient
Group
DICE
FUND
WITCH


Carbon Tax, Temperature, 2100
0.99 Population Growth
1.00 Population Growth
1.01 Population Growth
1.01 Economic Growth
1.00 Economic Growth
0.99 Economic Growth
Temperature
Temperature Sensitivity Coefficient
Group
DICE
FUND
GCAM
IGSM
MERGE
WITCH

Appendix 4. Additional Tables and Graphs

Table A1. Overview of global integrated assessment models included in this study.

Model	Number of Economic Regions	Time Horizon	Variables Included	Key Characteristics	Selected References
DICE	1	2010-2300	1,2,3,5,6	Optimal growth model, endogenous GDP and temperature, exogenous population, SWF is CES with respect to consumption.	(Nordhaus and Sztorc 2014)
FUND	16	1950-3000	1,2,3,4,5,6,7	Multi-region, multi-gas, detailed damage functions, exogenous scenarios perturbed by model	(Anthoff and Tol 2010, 2013)
GCAM	14	2005-2095	1,2,3,4,5,7	Integrated energy-land-climate model with technology detail; exogenous population and GDP; endogenous energy resources, agriculture, and temperature; economic costs are calculated for producer and consumer surplus change	(Calvin and et al. 2011)
IGSM	16	2100	1,2,3,4,5,7	Full general circulation model linked to a multi sector-multi region general equilibrium model of the economy with explicit advanced technology options	(Chen et al. 2015, Sokolov et al. 2009, Webster et al. 2012)
MERGE	10	2100	1,2,3,4,5,7	Ramsey model coupled	(Blanford

				with energy process model, multiple regions, endogenous GDP and temperature, exogenous population	et al. 2014)
WITCH	13	2150	1,2,3,4,5,6,7	Optimal growth model, endogenous GDP and temperature, exogenous population, SWF is CES with respect to consumption.	(Bosetti et al. 2006)

Notes: SWF = social welfare function, CES = constant elasticity of substitution. For variables included the key is:
1 = GDP, population
2 = CO2 emissions, CO2 concentrations
3 = global temperature
4 = multiple regions
5 = mitigation
6 = damages
7 = non-CO2 GHGs

Results of Monte Carlo simulations for models and major variables
[All variables are 2100 except SCC, which is 2020]

model	variable	mean	sigma	0.1%ile	1%ile	5%ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	99.9%ile
DICE	CO2 Conc	921.6	230.6	206.6	384.6	542.3	626.0	765.8	921.7	1,077.3	1,216.9	1,300.5	1,457.1	1,635.8
FUND	CO2 Conc	877.8	239.2	143.3	320.6	484.2	571.0	716.8	877.7	1,039.0	1,184.6	1,271.7	1,435.1	1,618.3
GCAM	CO2 Conc	898.7	173.6	361.7	494.8	613.2	676.3	781.4	898.6	1,015.7	1,121.4	1,184.5	1,303.3	1,435.3
IGSM	CO2 Conc	813.1	99.0	507.2	582.3	650.2	686.2	746.4	813.0	879.7	940.0	976.2	1,043.5	1,119.2
MERGE	CO2 Conc	973.8	231.7	259.9	434.5	592.5	676.8	817.4	973.6	1,130.3	1,271.1	1,354.7	1,512.6	1,688.8
WITCH	CO2 Conc	845.0	104.4	522.8	603.3	673.5	711.4	774.5	845.0	915.4	978.8	1,016.9	1,088.1	1,169.4
DICE	Temp	3.703	0.946	1.197	1.753	2.266	2.552	3.049	3.640	4.288	4.932	5.354	6.232	7.400
FUND	Temp	3.507	0.565	1.800	2.213	2.587	2.788	3.125	3.502	3.884	4.234	4.445	4.849	5.315
GCAM	Temp	3.666	0.926	1.197	1.744	2.254	2.535	3.025	3.607	4.239	4.867	5.279	6.131	7.240
IGSM	Temp	3.485	0.637	1.562	2.030	2.448	2.674	3.053	3.478	3.909	4.304	4.543	4.999	5.526
MERGE	Temp	3.942	0.907	1.572	2.091	2.574	2.842	3.313	3.877	4.498	5.122	5.533	6.388	7.523
WITCH	Temp	3.329	0.754	1.550	1.898	2.241	2.440	2.798	3.250	3.772	4.316	4.682	5.468	6.524
DICE	Output	675.62	443.06	(695.55)	(356.51)	(53.66)	107.67	376.78	675.69	974.64	1,243.56	1,404.09	1,704.52	2,039.22
FUND	Output	445.54	316.87	(529.01)	(291.21)	(75.58)	39.01	232.06	445.39	659.16	851.68	967.49	1,184.00	1,424.52
GCAM	Output	611.87	429.64	(717.06)	(384.65)	(94.97)	60.99	321.45	611.99	901.48	1,162.94	1,318.88	1,612.57	1,937.46
IGSM	Output	527.50	358.42	(577.88)	(308.26)	(61.58)	68.86	286.27	527.23	769.14	986.75	1,117.01	1,361.64	1,641.10
MERGE	Output	623.91	450.73	(769.54)	(422.22)	(117.14)	45.30	319.45	624.03	928.45	1,201.57	1,364.55	1,668.47	2,024.60
WITCH	Output	621.29	445.47	(751.02)	(413.24)	(110.92)	50.46	320.63	620.37	920.97	1,192.52	1,354.97	1,659.49	2,005.51
DICE	CO2 emission	123.21	67.49	(84.76)	(33.71)	12.11	36.64	77.70	123.24	168.70	209.70	234.20	280.21	331.84
FUND	CO2 emission	130.02	92.34	(155.29)	(84.99)	(22.13)	11.57	67.83	130.05	192.28	248.34	281.86	344.55	415.44
GCAM	CO2 emission	104.23	40.78	(21.98)	9.39	37.17	51.98	76.68	104.22	131.79	156.46	171.31	198.93	229.75
IGSM	CO2 emission	71.75	22.55	2.23	19.28	34.68	42.85	56.55	71.76	86.98	100.65	108.80	124.15	141.70
MERGE	CO2 emission	155.11	87.86	(116.08)	(49.75)	10.61	42.51	95.87	155.11	214.44	267.69	299.62	358.90	425.84
WITCH	CO2 emission	91.46	26.38	10.25	30.11	48.03	57.69	73.68	91.44	109.27	125.32	134.87	152.67	172.73
DICE	Pop	12,317	2,421	4,826	6,675	8,334	9,214	10,682	12,323	13,949	15,417	16,298	17,948	19,814
FUND	Pop	12,473	2,445	4,915	6,805	8,455	9,344	10,822	12,471	14,123	15,610	16,496	18,157	20,044
GCAM	Pop	10,597	2,080	4,166	5,756	7,178	7,933	9,192	10,598	12,001	13,260	14,019	15,439	17,012
IGSM	Pop	13,128	2,569	5,193	7,167	8,906	9,837	11,396	13,126	14,859	16,423	17,352	19,126	21,085
MERGE	Pop	12,195	2,391	4,792	6,632	8,259	9,125	10,583	12,196	13,806	15,262	16,127	17,756	19,574
WITCH	Pop	12,166	2,377	4,792	6,640	8,267	9,120	10,563	12,166	13,766	15,213	16,073	17,700	19,513
DICE	Rad Forc	7.06	1.35	2.90	3.93	4.85	5.34	6.16	7.06	7.97	8.79	9.28	10.19	11.22
FUND	Rad Forc	7.34	1.39	3.06	4.10	5.05	5.56	6.40	7.34	8.28	9.12	9.62	10.58	11.64
GCAM	Rad Forc	7.55	1.26	3.66	4.61	5.47	5.93	6.70	7.55	8.41	9.17	9.63	10.50	11.45
IGSM	Rad Forc	7.88	1.51	3.22	4.36	5.39	5.95	6.87	7.89	8.90	9.81	10.36	11.39	12.55
MERGE	Rad Forc	7.51	1.29	3.52	4.51	5.39	5.86	6.64	7.51	8.38	9.17	9.64	10.52	11.50
WITCH	Rad Forc	7.08	0.64	5.12	5.60	6.03	6.26	6.65	7.08	7.51	7.90	8.13	8.56	9.05

model	variable	mean	sigma	0.1 %ile	1 %ile	5 %ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	99.9%ile
DICE	SCC	21.6137	13.8369	1.7352	3.6969	6.0437	7.7603	11.8255	18.4038	27.8106	39.3834	48.0585	68.9073	101.7279
FUND	SCC	2.5469	1.8042	(0.8916)	(0.2173)	0.4211	0.7701	1.3404	2.1062	3.3711	4.9283	6.0640	8.5806	12.0513
WITCH	SCC	50.9695	16.7251	15.8230	21.4223	27.6464	31.5554	38.9603	48.7988	60.6450	73.1721	81.6266	99.4283	123.1576

model	variable	mean	sigma	0.1 %ile	1 %ile	5 %ile	10%ile	25%ile	50%ile	75%ile	90%ile	95%ile	99%ile	99.9%ile
DICE	Damages	34.54	47.03	(12.53)	(7.96)	(4.02)	(1.97)	1.99	17.32	50.20	95.95	130.41	210.22	320.68
FUND	Damages	3.55	5.03	(2.79)	(1.92)	(1.05)	(0.54)	0.36	1.79	5.18	10.05	13.77	22.30	34.58
WITCH	Damages	36.23	37.50	(3.97)	1.36	4.20	5.37	9.17	23.21	50.31	85.93	112.38	171.89	255.20

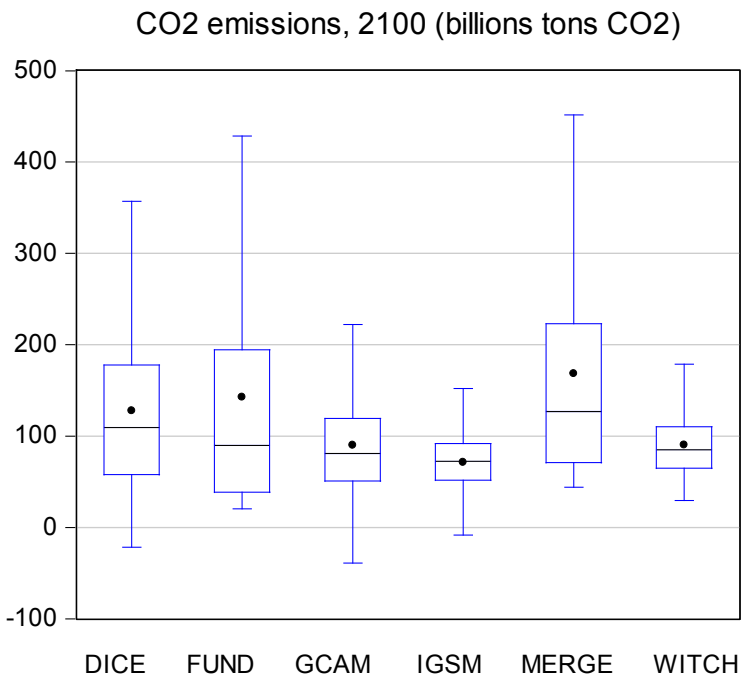


Figure for box plots for CO_2 emissions, 2100. For discussion of box plots, see Figure 7.

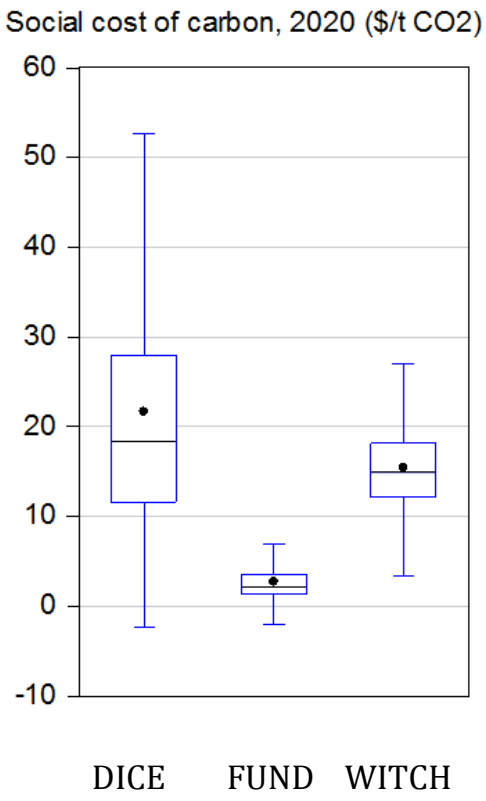


Figure for box plots for social cost of carbon, 2020. For discussion of box plots, see Figure 7.

Model	Variable	1-R2	Standard error regression (SER)	Mean variable	SER/Mean
DICE	CO2 Conc	0.00136	10.9	851.4	0.01277
FUND	CO2 Conc	0.00857	28.6	796.7	0.03587
GCAM	CO2 Conc	0.00132	11.5	825.7	0.01398
IGSM	CO2 Conc	0.00020	1.5	811.5	0.00187
MERGE	CO2 Conc	0.00468	19.3	905.6	0.02136
WITCH	CO2 Conc	0.00052	3.8	826.1	0.00464
DICE	Temp	0.00073	0.05	3.82	0.01428
FUND	Temp	0.00529	0.06	3.61	0.01715
GCAM	Temp	0.00423	0.15	3.66	0.04043
IGSM	Temp	0.00283	0.05	3.57	0.01295
MERGE	Temp	0.00388	0.14	4.25	0.03239
WITCH	Temp	0.00196	0.08	3.76	0.02258
DICE	ln output	0.00012	85.2	489.4	0.00969
FUND	ln output	0.00000	58.0	305.6	0.00042
GCAM	ln output	0.00000	76.3	425.6	0.00012
IGSM	ln output	0.00019	18.6	413.1	0.01132
MERGE	ln output	0.00001	78.3	439.7	0.00268
WITCH	ln output	0.00000	73.8	436.0	0.00139

Note: For ln equation, last column is logarithmic standard error of regression.

Estimates from surface response functions by variable and model.

Specification	DICE	FUND	GCAM	IGSM	MERGE	WITCH
L	0.06442	0.20137	0.11755	0.06091	0.11929	0.06312
LQ	0.05364	0.15166	0.06487	0.00675	0.09230	0.02608
LI	0.06146	0.06813	0.06048	0.06091	0.06094	0.06081
LQI	0.00706	0.01842	0.00621	0.00283	0.00733	0.00196
LQI++	0.00687	0.01842	0.00621	0.00139	0.00733	0.00138

Goodness of fit of worst fitting LQI variable by model.

Table shows the residual variance (1-R2) for the worst fitting of the equations. For example, in the LQI specification, the worst SRF for the DICE model is the equation for population, which has a residual variance of 0.00706. For the MERGE model, the worst equation is for CO_2 emissions. Note as well that the only two models for which the worst equation has a significant reduction in residual variation from LQI to LQI++ are the IGSM and WITCH models.

References

Acemoglu, D., S. Johnson, and J. Robinson. 2005. "Institutions as a Fundamental Cause of Long-run Growth." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven Durlauf. North-Holland.

Anthoff, D., and R. Tol (2010). "On International Equity Weights and National Decision Making on Climate Change." *Journal of Environmental Economics and Management* 60(1): 14-20.

Anthoff, D., and R. Tol (2013). "The Uncertainty About the Social Cost of Carbon: A Decomposition Analysis Using FUND." *Climatic Change* 117(3): 515-530.

Armstrong, J. Scott. (2001). "Combining forecasts." *Principles of forecasting.* Springer US, 417-439.

Baker, E. (2005). "Uncertainty and Learning in a Strategic Environment: Global Climate Change." *Resource and Energy Economics* 27(1): 19-40.

Batchelor, Roy, and Pami Dua (1995). "Forecaster diversity and the benefits of combining forecasts." Management Science 41.1 (1995): 68-75.

Blanford, G., J. Merrick, R. Richels, and R. Steven (2014). "Trade-offs Between Mitigation Costs and Temperature Change." *Climatic Change* 123(3-4): 527-541.

Bosetti, V., C. Carraro, M. Galeotti, E. Massetti, and M. Tavoni (2006). "WITCH: A World Induced Technical Change Hybrid Model." *Energy Journal* 27(SI2): 13-37.

Bosetti, V., C. Carraro, E. Massetti, and M. Tavoni. 2014. *Climate Change Mitigation, Technological Innovation and Adaptation*: Edward Elgar Publishers.

Brynjolfsson, E., and A. McAfee. 2012. *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*: Digital Frontier Press.

Calvin, K., and et al. 2011. GCAM Wiki Documentation. http://wiki.umd.edu/gcam/index.php?title=Main_Page. College Park, MD: Joint Global Change Research Institute.

CBO. 2005. Uncertainty in Analyzing Climate Change: Policy Implications. Washington, DC: Congressional Budget Office.

Chen, Y.-H., S. Paltsev, J. Reilly, J.F. Morris, and M.H. Babiker. 2015. The MIT EPPA6 Model: Economic Growth, Energy Use, and Food Consumption, MIT Joint Program Report Number 278. Cambridge, MA.

Clarke, L., and J. Weyant (2009). "Introduction to the EMF 22 Special Issue on Climate Change Control Scenarios." *Energy Economics* 31(2): S63.

Clemen, Robert T., and Robert L. Winkler (1999). "Combining probability distributions from experts in risk analysis." Risk analysis 19.2: 187-203.

Clements, M., and D. Hendry. 1998. *Forecasting Economic Time Series*. Cambridge, UK: Cambridge University Press.

Clements, M., and D. Hendry. 1999. *Forecasting Non-stationary Economic Time Series*. Cambridge, MA: MIT Press.

de Finetti, B. (1937). "La prevision: Ses lois logiques, ses sources subjectives." *Annales de l'Institut Henri Poincaré* 7: 1-68.

Edmonds, J., and J. Reilly (1983a). "Global Energy and CO2 to the Year 2050." *Energy Journal* 4(3): 21-47.

Edmonds, J., and J. Reilly (1983b). "Global Energy Production and Use to the Year 2050." *Energy* 8(6): 419-432.

Edmonds, J., and J. Reilly (1983c). "A Long-term Global Energy-economic Model of Carbon Dioxide Release From Fossil Fuel Use." *Energy Economics* 5(2): 74-88.

Ericsson, N. 2001. Forecast Uncertainty in Economic Modeling. Washington, DC: Board of Governors of the Federal Reserve System International Finance Discussion Papers.

Fernald, J., and C. Jones. 2014. The Future of U.S. Economic Growth. Cambridge, MA: National Bureau of Economic Research Working Paper 19830

Freeman, R. 2010. "What Does Global Expansion of Higher Education Mean for the United States?" In *American Universities in a Global Market*, 373-404. University of Chicago Press.

Gordon, R. 2012. Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds. Cambridge, MA: National Bureau of Economic Research Working Paper 18315.

Greenstone, M., E. Kopits, and A. Wolverton (2013). "Developing a Social Cost of Carbon for US Regulatory Analysis: A Methodology and Interpretation." *Review of Environmental Economics and Policy* 7(1): 23-46.

Hammersley, J.M., and D.C. Handscomb. 1964. *Monte Carlo Methods*. New York: John Wiley and Sons.

Hope, C. (2006). "The Marginal Impact of CO2 from PAGE2002: An Integrated Assessment Model Incorporating the IPCC's Five Reasons for Concern." *Integrated Assessment* 6(19-56).

IAWG. 2010. Technical Support Document: Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. Washington, DC: Interagency Working Group on the Social Cost of Carbon.

IAWG. 2013. Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. Washington, DC: Interagency Working Group on the Social Cost of Carbon.

InterAcademy Council. 2010. Climate Change Assessments: Review of the Processes and Procedures of the IPCC, 2010, Harold Shapiro, chair.

IPCC. 2014. *Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY: Cambridge University Press.

Knutti, R., and G. Hegerl (2008). "The Equilibrium Sensitivity of the Earth's Temperature to Radiation Changes." *Nature Geoscience* 1: 735-743.

Kriegler, E., N. Peterman, V. Krey, V.J. Schwanitz, G. Luderer, S. Ashina, V. Bosetti, J. Eom, A. Kitous, A. Mejean, L. Paroussos, F. Sano, H. Turton, C. Wilson, and D. Van Vuuren (2015). "Diagnostic Indicators for Integrated Assessment Models of Climate Change." *Technological Forecasting and Social Change* 90(A): 45-61.

Lemoine, D., and H. McJeon (2013). "Trapped Between Two Tails: Trading Off Scientific Uncertainties via Climate Targets." *Environmental Research Letters* 8: 1-10.

Lenton, T., H. Held, E. Kriegler, J. Hall, W. Lucht, S. Rahmstorf, and H.J. Schellnhuber (2008). "Tipping Elements in the Earth's Climate System." *Proceedings of the National Academy of Sciences* 105(6): 1786-1793.

Lutz, W., ed. 1996. *The Future Population of the World: What Can We Assume Today?* London: Earthscan Publication Ltd.

Lutz, W., W. Butz, and S. KC. 2014. *World Population and Human Capital in the Twenty-First Century*. Oxford, UK: Oxford University Press.

Lutz, W., W. Sanderson, and S. Scherbov. 1998. "Expert-based Probabilistic Projections." In *Frontiers of Population Forecasting*, edited by Wolfgang Lutz, J.W. Vaupel and D.A. Ahlburg, 139-155. Population and Development Review.

Lutz, W., W. Sanderson, and S. Scherbov. *IIASA's 2007 Probabilistic World Population Projections, IIASA World Population Program Online Data Base of Results* 2008. Available from http://www.iiasa.ac.at/web/home/research/researchPrograms/WorldPopulation/Reaging/2007_update_prob_world_pop_proj.html.

Manne, A., R. Mendelsohn, and R. Richels (1999). "MERGE: A Model for Evaluating Regional and Global Effects of Greenhouse Gas Reduction Policies." *Energy Policy* 23(1): 17-34.

Meinshausen, M., S.C. Raper, and T. Wigley (2011). "Emulating Coupled Atmosphere-Ocean and Carbon Cycle Models with a Simpler Model, MAGICC6 - Part I:

Model Description and Calibration." *Atmospheric Chemistry and Physics* 11: 1417-1456.

Nordhaus, W. 2008. *A Question of Balance: Weighing the Options on Global Warming Policies*. New Haven, CT: Yale University Press.

Nordhaus, W., and D. Popp (1997). "What is the Value of Scientific Knowledge? An Application to Global Warming Using the PRICE Model." *Energy Journal* 18(1): 1-45.

Nordhaus, W., and P. Sztorc. 2014. DICE 2013: Introduction and User's Manual. New Haven, CT: Yale University.

NRC. 2000. *Beyond Six Billion: Forecasting the World's Population*. Washington, DC: National Academy Press.

O'Neill, B., D. Balk, M. Brickman, and M. Ezra (2001). "A Guide to Global Population Projections." *Demographic Research* 4(8): 203-288.

Olsen, R., R. Sriver, M. Goes, N. Urban, D. Matthews, M. Haran, and K. Keller (2012). "A Climate Sensitivity Estimate Using Bayesian Fusion of Instrumental Observations and an Earth System Model." *Geophysical Research Letters* 117(D04103): 1-11.

Peck, S., and T. Teisberg (1993). "Global Warming Uncertainties and the Value of Information: An Analysis Using CETA." *Resource and Energy Economics* 15(1): 71-97.

Pizer, W. (1999). "Optimal Choice of Climate Change Policy in the Presence of Uncertainty." *Resource and Energy Economics* 21(3-4): 255-287.

Pizer, W., M. Adler, J. Aldy, D. Anthoff, M. Cropper, K. Gillingham, M. Greenstone, B. Murray, R. Newell, R. Richels, A. Rowell, S. Waldhoff, and J. Wiener (2014). "Using and Improving the Social Cost of Carbon." *Science* 346(6214): 1189-1190.

Ramsey, F. 1931. "Truth and Probability." In *The Foundations of Mathematics and Other Logical Essays*, edited by Richard Bevan Braithwaite, 156-198. London, UK: Kegan, Paul, Trench, Trubner and Company.

Reilly, J., J. Edmonds, R. Gardner, A. Brenkert (1987) "Monte Carlo Analysis of the IEA/ORAU Energy/Carbon Emissions Model." *Energy Journal* 8:1-29.

Revesz, R., P. Howard, K. Arrow, L. Goulder, R. Kopp, M. Livermore, M. Oppenheimer, and T. Sterner (2014). "Global Warming: Improve Economic Models of Climate Change." *Nature* 508(7495): 173-175.

Robinson, A., R. Calov, and A. Ganopolski (2012). "Multistability and Critical Thresholds of the Greenland Ice Sheet." *Nature Climate Change* 2: 429-432.

Rytgaard, Mette (1990). "Estimation in the Pareto distribution*." Astin Bulletin* 20.02: 201-216.

Savage, L. 1954. *The Foundations of Statistics*. New York: John Wiley and Sons.

Schuster, Eugene F. (1984). "Classification of Probability Laws by Tail Behavior," *Journal of the American Statistical Association,* Vol. 79, No. 388: 936-939.

Sokolov, A., P.H. Stone, C. Forest, R. Prinn, M. Sarofim, M. Webster, S. Paltsev, A. Schlosser, D. Kicklighter, S. Dutkiewicz, J. Reilly, C. Wang, B. Felzer, J. Melillo, and H. Jacoby (2009). "Probabilitistic Forecast for 21st Century Climate Based on Uncertainties in Emissions (without Policy) and Climate Parameters." *Journal of Climate* 22(19): 5175-5204.

Tol, Richard (1997) "On the Optimal Control of Carbon Dioxide Emissions - An Application of FUND." *Environmental Modelling and Assessment*, 2: 151-163.

US Interagency Working Group. 2013. Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. Washington, DC: Executive Office of the President.

van Vuuren, D., B. de Vries, A. Beusen, and P. Heuberger (2008). "Conditional Probabilistic Estimates of 21st Century Greenhouse Gas Emissions Based on the Storylines of the IPCC-SRES Scenarios." *Global Environmental Change*, 18: 635-654.

Weaver, A., M. Eby, E. Wiebe, C. Bitz, P. Duffy, T. Ewen, A. Fanning, M. Holland, A. MacFadyen, D. Matthews, K. Meissner, O. Saenko, A. Schmittner, H. Wang, and M. Yoshimori (2001). "The UVic Earth System Climate Model: Model Description, Climatology, and Applications to Past, Present and Future Climates." *Atmosphere-Ocean* 39(4): 361-428.

Webster, M. (2002). "The Curious Role of Learning: Should We Wait for More Data?" *Energy Journal* 23(2): 97-119.

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

Webster, M., A. Sokolov, J. Reilly, C. Forest, S. Paltsev, A. Schlosser, C. Wang, D. Kicklighter, M. Sarofim, J. Melillo, R. Prinn, and H. Jacoby (2012). "Analysis of Climate Policy Targets Under Uncertainty." *Climatic Change* 112(3-4): 569-583.

Weitzman, M. (2009). "On Modeling and Interpreting the Economics of Catastrophic Climate Change." *Review of Economics and Statistics* 91(1): 1-19.