WEB APPENDIX

Modeling Uncertainty in Integrated Assessment of Climate Change: A Multi-Model Comparison

By KENNETH GILLINGHAM, WILLIAM NORDHAUS, DAVID ANTHOFF, GEOFFREY BLANFORD, VALENTINA BOSETTI, PETER CHRISTENSEN, HAEWON MCJEON, AND JOHN REILLY[1]

Appendix A. 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). The data underlying these pdfs is compiled for the IPCC AR5 and were shared with us. For each, we examined both the raw pdfs and fitted a log-normal pdf to each based on minimizing the sum of squared deviations. We found that the fit is extremely good for all of them and for most, it is nearly impossible to tell the difference between the log-normal fits and the raw data pdfs. Figure A1 illustrates the log-normal fits to each of these distributions.

[1] * Corresponding Authors: Kenneth Gillingham and William Nordhaus. Gillingham: Yale University, 195 Prospect Street New Haven, CT 06511 (kenneth.gillingham@yale.edu); Nordhaus: Yale University, 28 Hillhouse Avenue, New Haven, CT 06511 (william.nordhaus@yale.edu); Anthoff: UC Berkeley; Blanford: Electric Power Research Institute; Bosetti: Bocconi University; Christensen: University of Illinois Urbana-Champaign; McJeon: Joint Global Change Research Institute; Reilly: MIT. The authors are grateful for comments from many colleagues over the course of this project. These include individual scholars as well as those at seminars and workshops at Yale University, the University of California, Berkeley, and the Snowmass climate-change meetings organized under the aegis of the Energy Modeling Forum. 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. Bosetti acknowledges funding from the European Research Council 336703 – project RISICO. The Stanford Energy Modeling Forum has provided support through its Snowmass summer workshops.

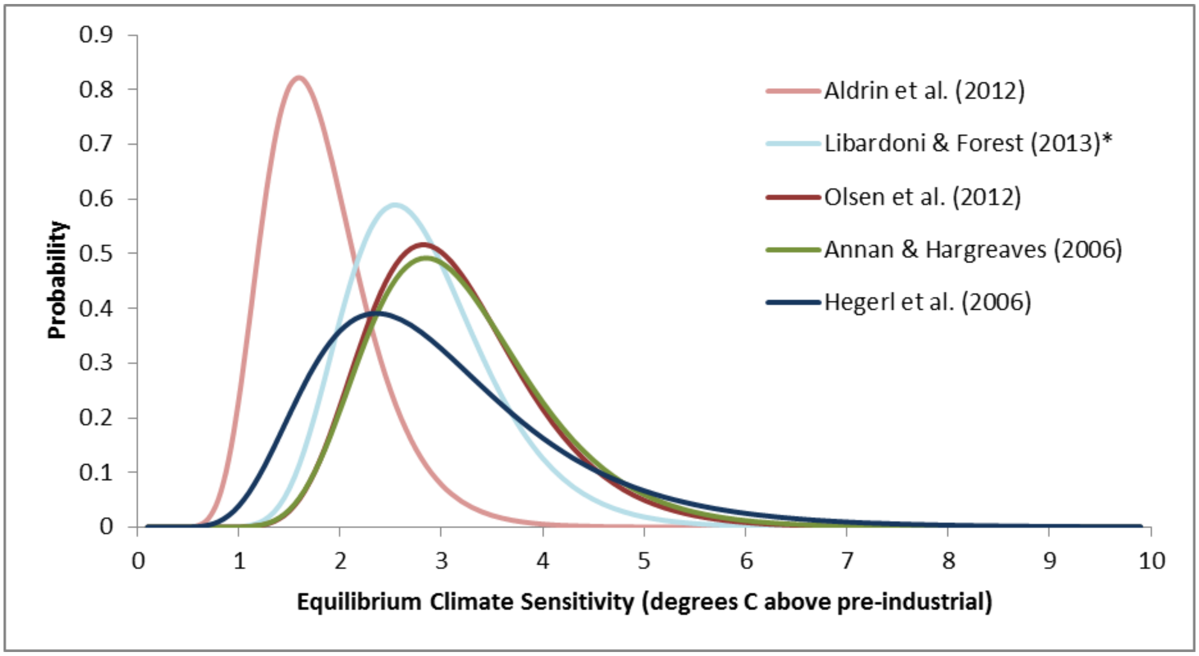

Figure A1. Log-normal distributions fit to the probability density functions cited in the IPCC AR5. The distribution shown here is from the updated Libardoni & Forest (2013) figures.

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 any two 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. Figure A2 shows the original and our fitted distributions.

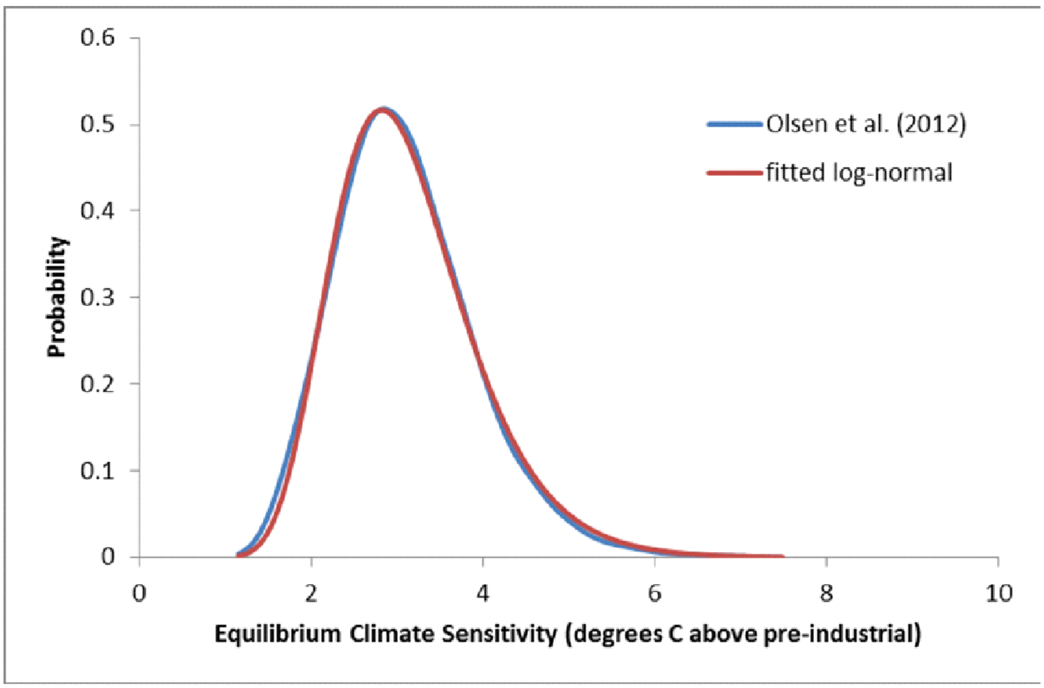

Figure A2. The Olsen et al. (2012) probability density function along with the fitted log-normal distribution used in our analysis.

Appendix B. Detailed description of the six models used in the MUP study

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 long-term global integrated assessment model originally developed by Edmonds and Reilly (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 2005-2100. The full documentation of the model is available at a GCAM wiki (Calvin and et al. 2011). GCAM is an open-source model, 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. (2016). 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.

Appendix C. Details on the procedures in tracks I and II

This appendix provides more detail on the track I and II procedures described in the main text. Begin with the first track. Each model has a baseline run with base values for each of the uncertain parameters. Denote the base parameter values as $(u_{m,1}^b, u_{m,2}^b, u_{m,3}^b)$. 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 $\Delta^G = (\Delta_{1,1,1}, \Delta_{1,1,2}, \ldots, \Delta_{5,5,5})$. 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 $\Delta_{1,1,1} = (-0.014, -0.02, -2)$. Then that calibration run would calculate the outcomes for $Y^m = H^m(z, \alpha, u_{m,1}^b - .014, u_{m,2}^b - 0.02, u_{m,3}^b - 2)$, where again $u_{m,k}^b$ 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 $Y^m = H^m(z, \alpha, u_{m,1}^b, u_{m,2}^b, u_{m,3}^b)$, 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) $$Y^m = R^m(u_1 - u_{m,1}^b, u_2 - u_{m,2}^b, u_3 - u_{m,3}^b) = R^m(u_{m,1}, u_{m,2}, u_{m,3})$$

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 distributions are calculated by Monte Carlo methods, for a sample size of *N*:

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

The notation here is that ${}^n u_{m,k}$ is the nth draw of random variable u_k in the Monte Carlo experiment for model *m*. 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 $\tilde{Y}^m$. 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).

Appendix D. 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. The runs are labelled as “Base” for completeness because another set of runs with carbon taxes (AMPERE) were also made but are not presented in this study.

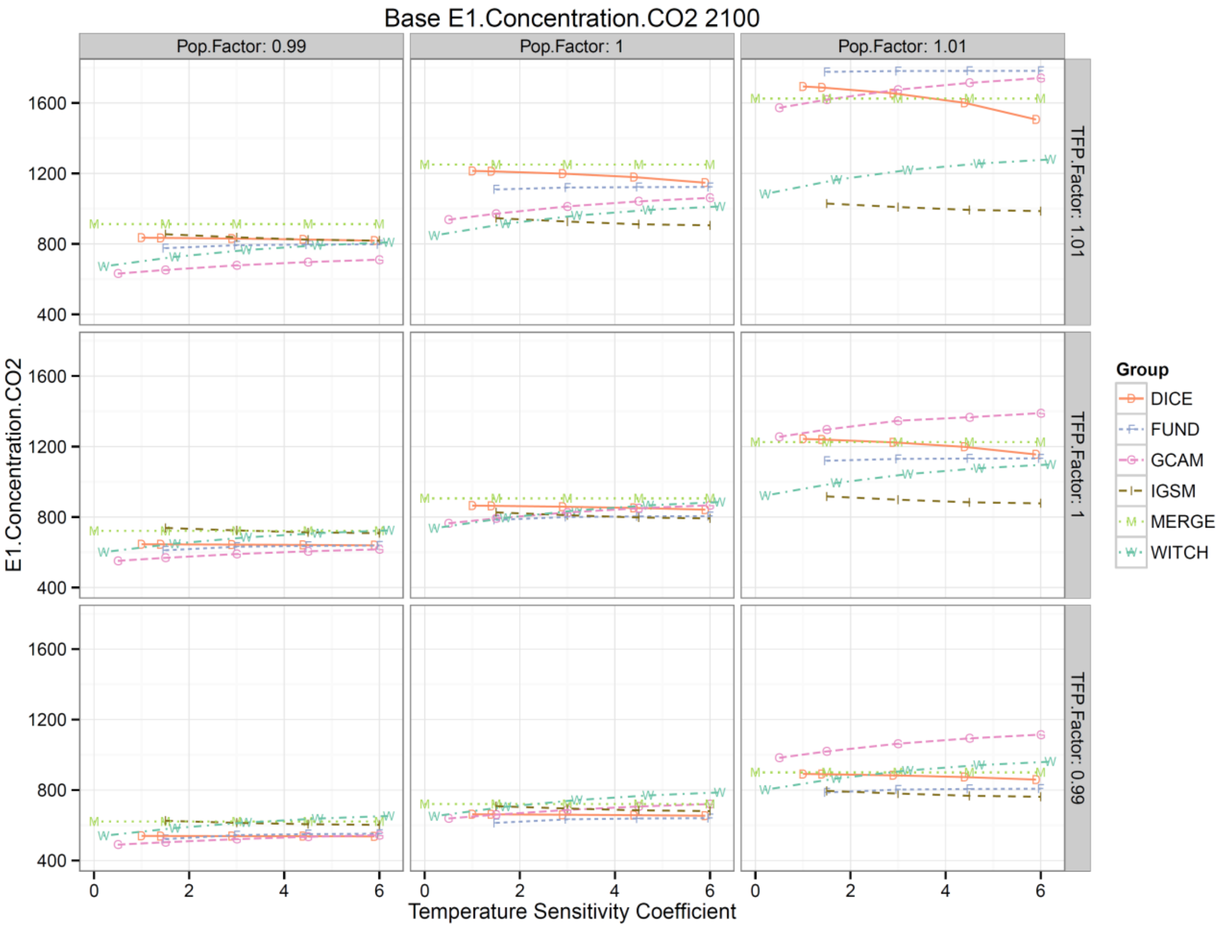

Figure D1. Lattice diagram for CO_2 concentrations in Base Case

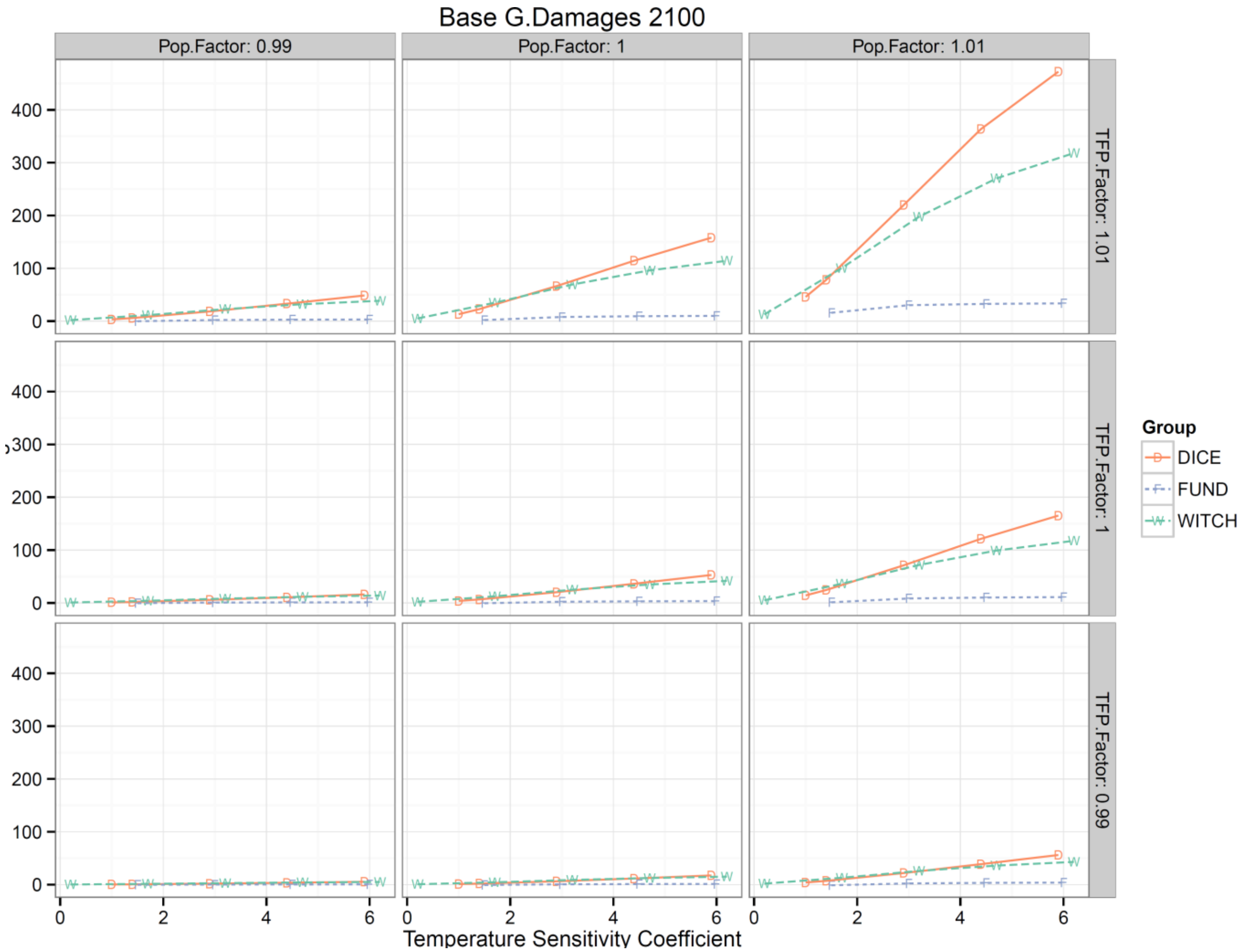

Figure D2. Lattice diagram for damages in Base Case

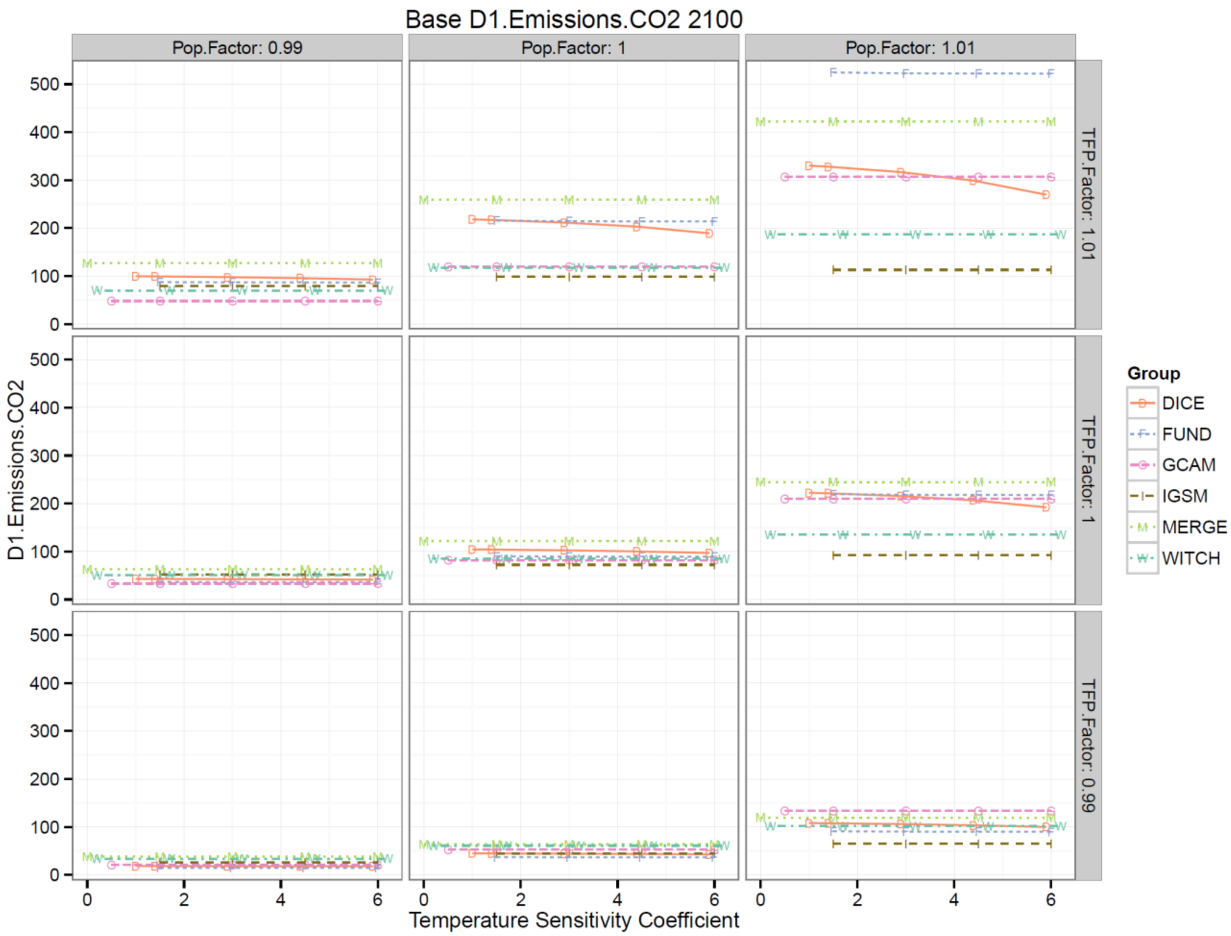

Figure D3. Lattice diagram for CO_2 emissions in Base Case

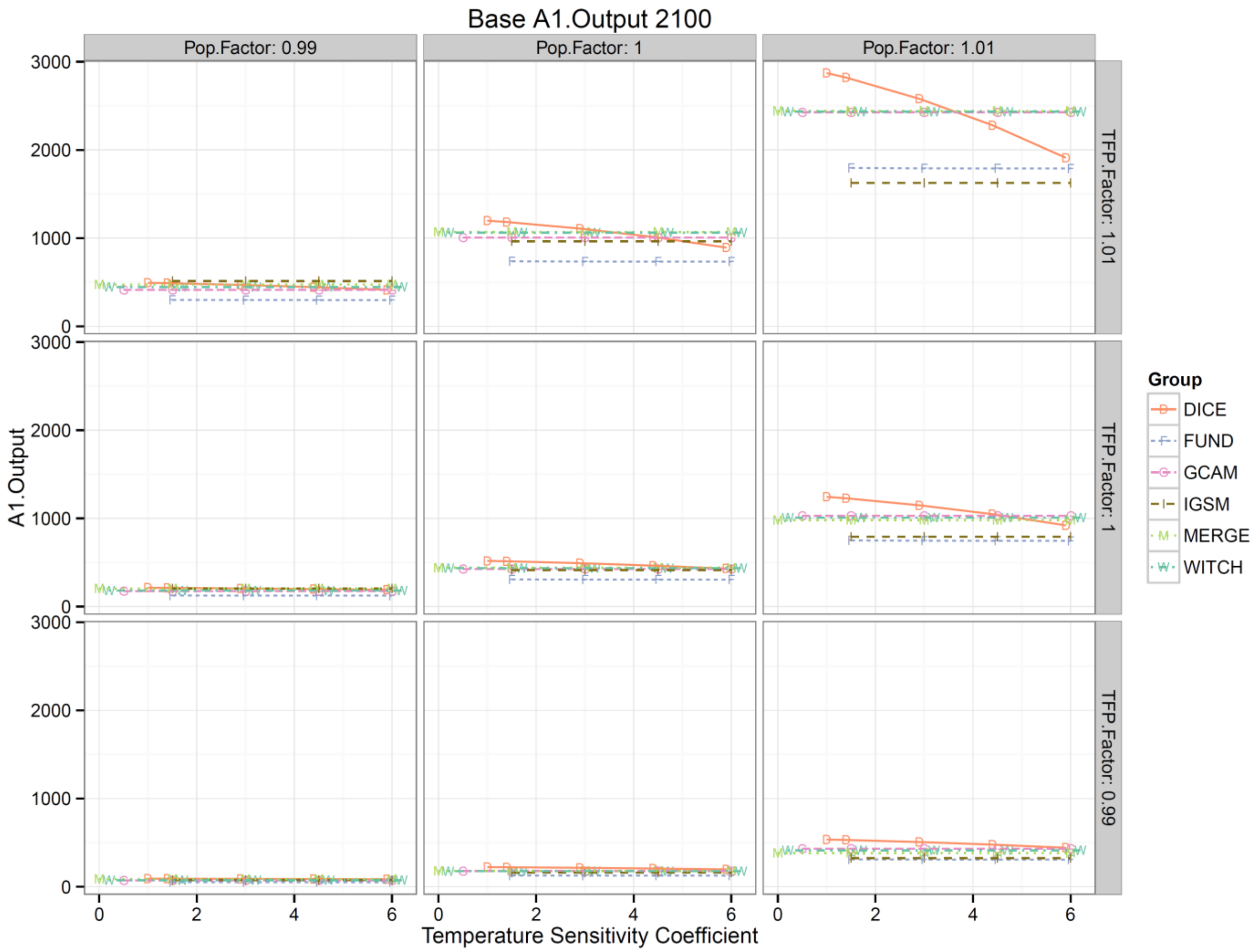

Figure D4. Lattice diagram for output in Base Case

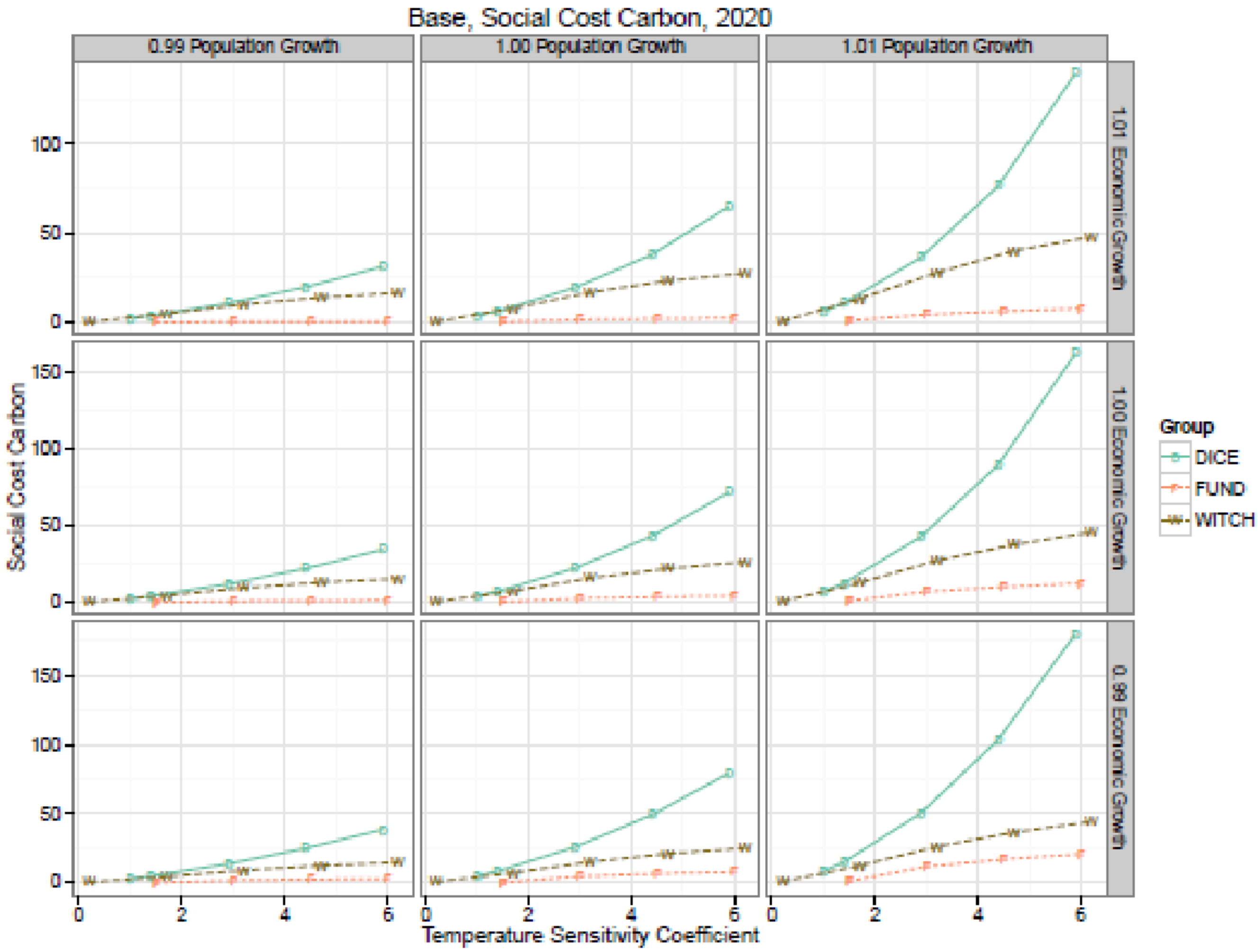

Figure D5. Lattice diagram for social cost of carbon in Base Case

Appendix E. Alternative Specifications of the Surface Response Function

We have focused on the LQI specification because of its simplicity, robustness, and high accuracy. However, we tested alternative specifications to determine whether the results changed significantly. For these tests we looked at the R^2 of the SRF equations fitted to the 125 observations for each model and variable; and the estimated standard deviation of the target variables (temperature, output, etc.) for each model.

At the simpler extreme of a linear specification, the results were indeed different. The R^2 of the SRF equation was much lower with the linear specification. The estimated standard deviation of the target variable on average was 3.1% lower with the linear specification compared to the LQI, with a standard deviation of 14%. This is an unacceptable difference.

When we tested a LQI++, with three added polynomial terms,[2] the SRF equations fit was essentially identical, with $R^2 = 0.9951$ for the LQI++ and 0.9948 for the LQI. The estimated standard deviations of the target variables were on average 0.2% lower with the LQI++ specification compared to the LQI, with a standard deviation of the ratio of 1.2%. In other words, the estimated uncertainties were virtually identical in the more complex specification. Figure A5 shows a histogram of ratios of the uncertainty of the LQI++ to LQI for 36 different models and variables.

The largest change between the LQI and the LQI++ was the uncertainty for temperature 2100 for the MERGE model, which decreased from 1.05 °C in the LQI specification to 1.00 °C in the LQI++ specification. The uncertainty for temperature 2100 for the DICE model, as another example, decreased from 1.114 °C in the LQI to 1.103 °C in the LQI++ specification.

We also examined higher order polynomials and found very little difference from the LQI++ specification. We also began to be concerned about over-fitting. The conclusion is that the LQI specification provides an accurate SRF to estimate the uncertainties in the two-step procedure.

[2] We tested specifications with more polynomial terms, but the data matrix was often singular. The specification which was able to run with all models and variables added only three additional non-linear terms. With this specification, 14 of 36 equations had R2 = 1.

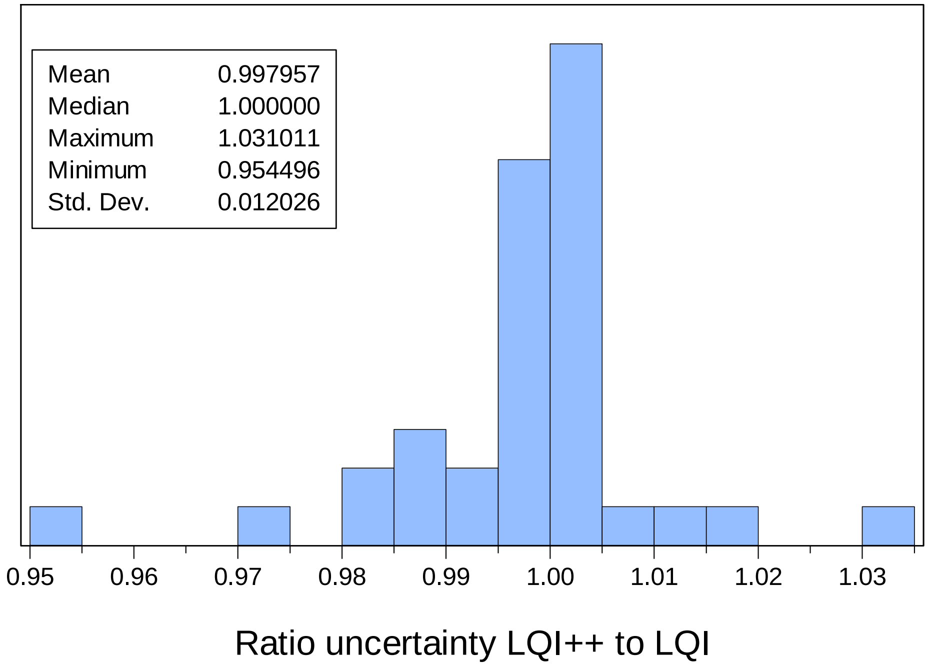

Figure E1. The histogram shows the ratio of the estimated uncertainty for major variables (output, temperature, etc.) under the two alternative specifications of the surface response function.

Appendix F. Additional Tables and Graphs

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	939	318	530	541	564	598	697	870	1,105	1,371	1,555	1,954	2,464
FUND	CO2 Conc	907	354	567	574	583	592	642	793	1,055	1,382	1,619	2,152	2,861
GCAM	CO2 Conc	861	222	472	508	563	606	696	826	988	1,160	1,276	1,518	1,824
IGSM	CO2 Conc	811	131	408	506	596	643	722	811	899	979	1,026	1,117	1,216
MERGE	CO2 Conc	998	325	633	637	645	666	750	917	1,159	1,440	1,637	2,070	2,630
WITCH	CO2 Conc	854	134	591	622	666	696	756	837	934	1,034	1,100	1,237	1,406
DICE	Temp	3.88	1.10	1.53	1.90	2.32	2.58	3.09	3.75	4.53	5.35	5.89	7.02	8.50
FUND	Temp	3.72	0.77	1.93	2.27	2.62	2.82	3.18	3.63	4.17	4.74	5.12	5.92	6.94
GCAM	Temp	3.94	1.02	1.54	1.97	2.43	2.70	3.21	3.85	4.57	5.30	5.76	6.71	7.88
IGSM	Temp	3.60	0.81	1.26	1.80	2.30	2.58	3.04	3.58	4.13	4.65	4.97	5.59	6.29
MERGE	Temp	4.31	1.01	2.16	2.51	2.90	3.13	3.59	4.19	4.90	5.65	6.15	7.19	8.50
WITCH	Temp	3.75	0.73	1.86	2.24	2.62	2.84	3.24	3.72	4.23	4.71	5.00	5.56	6.20
DICE	Output	735	690	115	129	146	159	233	495	997	1,653	2,136	3,239	4,717
FUND	Output	489	500	72	73	77	84	133	309	667	1,147	1,507	2,332	3,448
GCAM	Output	670	673	104	106	111	120	187	428	913	1,559	2,041	3,145	4,639
IGSM	Output	599	538	112	113	117	128	199	422	819	1,322	1,686	2,510	3,608
MERGE	Output	706	716	113	114	118	128	196	446	957	1,648	2,165	3,354	4,953
WITCH	Output	699	706	109	113	119	129	196	444	948	1,629	2,138	3,306	4,880
DICE	Emissions	128	92	(7)	1	13	26	58	110	178	253	304	415	555
FUND	Emissions	143	146	21	21	22	24	39	90	195	335	439	680	1,004
GCAM	Emissions	90	53	(3)	10	22	31	51	81	120	161	190	250	327
IGSM	Emissions	71	30	(31)	(3)	20	32	52	73	92	109	118	135	153
MERGE	Emissions	169	130	47	48	49	52	71	127	223	343	430	625	883
WITCH	Emissions	91	35	31	35	44	51	65	85	110	137	155	193	240
DICE	Rad Forcings	7.1	1.8	3.0	3.7	4.5	5.0	5.8	7.0	8.2	9.5	10.3	11.9	13.9
FUND	Rad Forcings	7.4	1.9	4.7	4.8	5.1	5.3	6.0	7.1	8.5	10.0	11.1	13.3	16.1
GCAM	Rad Forcings	7.3	1.6	3.3	4.1	4.9	5.4	6.2	7.3	8.3	9.4	10.0	11.4	12.9
IGSM	Rad Forcings	7.8	2.0	1.8	3.3	4.6	5.3	6.4	7.8	9.2	10.4	11.2	12.6	14.3
MERGE	Rad Forcings	7.5	1.7	4.4	4.7	5.2	5.5	6.3	7.3	8.5	9.8	10.7	12.5	14.8
WITCH	Rad Forcings	7.1	0.8	5.0	5.4	5.9	6.1	6.6	7.1	7.7	8.2	8.5	9.2	10.0
DICE	Damages	47.0	66.7	-14.3	-8.5	-4.0	-1.8	2.7	21.6	65.8	131.6	182.7	304.6	475.5
FUND	Damages	5.1	7.3	-2.9	-1.9	-1.0	-0.5	0.5	2.4	7.1	14.3	20.0	33.5	52.6
WITCH	Damages	45.3	52.1	-5.7	1.4	4.3	5.5	9.4	25.8	61.8	112.7	151.3	241.7	366.7
DICE	SCC	21.9	15.3	1.0	3.0	5.3	7.0	11.2	18.2	28.4	41.2	51.0	75.0	114.4
FUND	SCC	2.8	2.2	-1.0	-0.3	0.4	0.8	1.4	2.1	3.6	5.6	7.1	10.4	15.2
WITCH	SCC	15.5	4.5	5.6	7.2	9.1	10.2	12.3	15.0	18.1	21.4	23.6	28.1	33.9

Table F1. Results of Monte Carlo simulations for models and major variables

[All variables are 2100 except SCC, which is 2020]

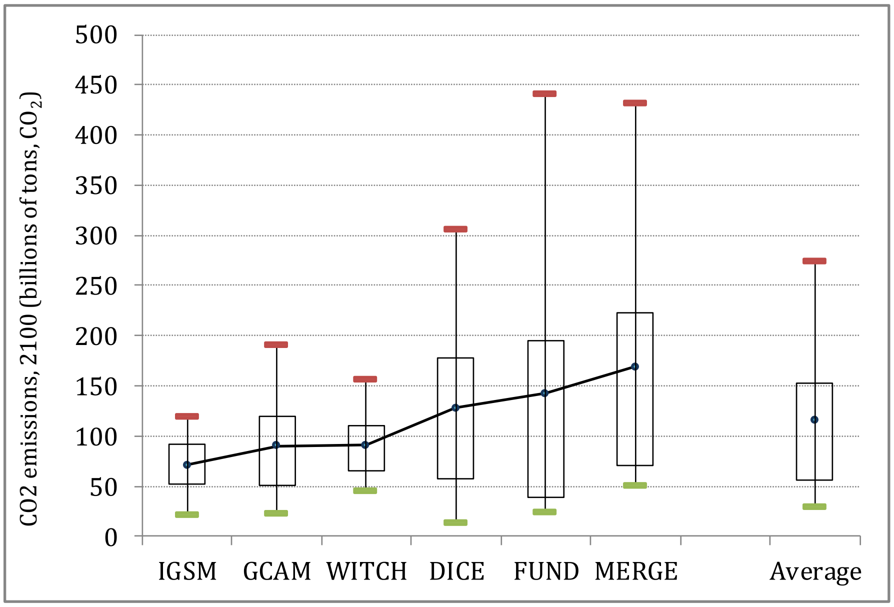

Figure F1. Box plots for CO_2 emissions, 2100.

For discussion of box plots, see Figure 5 in the main text.

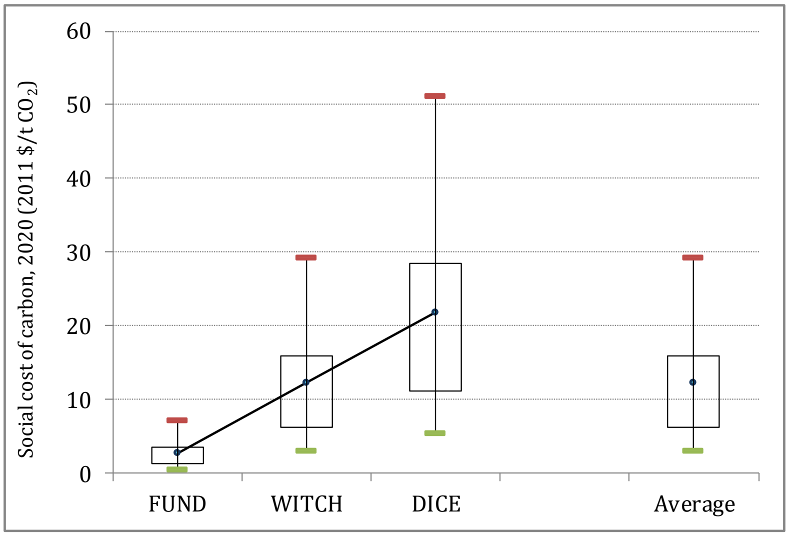

Figure F2. Box plots for social cost of carbon, 2020.

For discussion of box plots, see Figure 5 in the main text.

Model	Variable (2100)	1-R2	Standard error regression (SER)	Mean variable	SER/Mean
DICE	CO_2 Conc	0.00136	10.87	851.4	0.01277
FUND	CO_2 Conc	0.00857	28.58	796.7	0.03587
GCAM	CO_2 Conc	0.00132	11.55	825.7	0.01398
IGSM	CO_2 Conc	0.00020	1.52	811.5	0.00187
MERGE	CO_2 Conc	0.00468	19.34	905.6	0.02136
WITCH	CO_2 Conc	0.00052	3.83	826.1	0.00464
DICE	Temp	0.00072	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	0.00282	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	0.00969		(a)
FUND	ln(output)	0.00000	0.00042		(a)
GCAM	ln(output)	0.00000	0.00012		(a)
IGSM	ln(output)	0.00019	0.01132		(a)
MERGE	ln(output)	0.00001	0.00268		(a)
WITCH	ln(output)	0.00000	0.00139		(a)
DICE	Emissions	0.00706	7.15	101.9	0.07017
FUND	Emissions	0.01842	16.86	89.3	0.18886
GCAM	Emissions	0.00621	6.06	81.8	0.07405
IGSM	Emissions	0.00139	0.90	72.2	0.01245
MERGE	Emissions	0.00733	9.36	121.5	0.07702
WITCH	Emissions	0.00138	1.48	85.3	0.01732

(a) The equation is logarithmic, so the SER is approximately the ratio of the standard error of estimate to the mean.

Table F2. Estimates from surface response functions by variable and model.