ON THE PHYSICS OF THREE INTEGRATED ASSESSMENT MODELS

RAPHAEL CALEL AND DAVID A. STAINFORTH

We untangle differences in the physical assumptions embedded in three influential integrated assessment models. Separating the impact of physical and economic assumptions facilitates more interpretable model intercomparisons.

Integrated assessment models (IAMs) couple simple models of the climate and the economy in order to simulate the global economic impacts of climate change under different mitigation scenarios. IAMs are frequently used to inform domestic climate change policy and international negotiations (Clarke et al. 2014), and policy makers have relied heavily

AFFILIATIONS: Calel—McCourt School of Public Policy, Georgetown University, Washington, D.C., and Grantham Research Institute on Climate Change and the Environment, London School of Economics, London, United Kingdom; Stainforth—Grantham Research Institute on Climate Change and the Environment, and Centre for the Analysis of Time Series, London School of Economics, London, and Department of Physics, University of Warwick, Coventry, United Kingdom
CORRESPONDING AUTHOR: Raphael Calel, raphael.calel@georgetown.edu

on three IAMs in particular—Dynamic Integrated model of Climate and the Economy (DICE), The Climate Framework for Uncertainty, Negotiation and Distribution (FUND), and Policy Analysis for the Greenhouse Effect (PAGE)—when trying to balance the benefits and costs of climate action (Stern 2007; Watkiss and Hope 2011; Greenstone et al. 2013; Interagency Working Group 2010, 2015; EPA 2014; Hahn and Ritz 2015).

These IAMs incorporate greatly simplified representations of the climate system compared to the complicated global climate models (GCMs) used by many climate scientists. This allows them to solve a demanding set of problems—whether finding an optimal policy in the case of DICE, or evaluating climate policies over ensembles of many parameter values in the case of FUND and PAGE.[1] This simplicity could also allow these IAMs to serve as a bridge across disciplines, giving economists and scientists a common

---

[1] Another important class of IAMs is set up to find the least-cost way of achieving a prespecified target. This is, in some sense, a less demanding problem, and these models can therefore incorporate climate modules that are a great deal more complex. The IPCC's Fifth Assessment Report lists over 30 such IAMs in use (Clarke et al. 2014). Many of them use MAGICC to represent the physical climate (see online supplement).

framework for pinning down the sources of policy disagreements. But this is only possible if the models are transparent enough for both research communities to engage with them.

Economists have previously highlighted several economic modeling choices that can have very large effects on the policy recommendations of these IAMs, including the choice of discount rate (Dasgupta 2007; Nordhaus 2007; Weitzman 2007) and the curvature of the damage function (Ackerman et al. 2010; Weitzman 2012). These economic sources of disagreement are now the subject of exciting new research (Millner and Heal 2014; Burke et al. 2015b). Unfortunately, in failing to maintain clear links to the physical science literature, the climate components of these models have become opaque to the scientific community. Researchers have consequently had less success understanding the physical sources of disagreement between these IAMs. This makes it more difficult for economists and scientists to work together to ensure that the IAMs reflect the latest climate science. To draw attention to economically important areas of scientific disagreement, as they have already done for economic disagreements, these IAMs must be made more transparent to the broader scientific community.

The authors of these IAMs have already made commendable efforts to publish (Nordhaus 2013; Anthoff and Tol 2013) and document (Nordhaus and Sztorc 2013; Hope 2006, 2011; Anthoff and Tol 2014) their models, but even with these aids it is often prohibitively difficult for the broader community of researchers to penetrate their terminology, notation, and coding. This makes it difficult to understand why they produce different policy recommendations. In response, recent studies have begun comparing the outputs of these IAMs more systematically and comparing them collectively with models found in the scientific literature. Van Vuuren et al. (2011) and Rose et al. (2014, chapter 5) compare IAMs with Earth systems models and with an intermediate-complexity climate model [Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC); please see online supplement at http://dx.doi.org/10.1175/BAMS-D-16-0034.2], respectively, and also subject the models to computational experiments to explore how they respond to identical emissions, concentrations, and forcings.[2] While informative, these efforts have been confined to comparing

models based on their outputs, without engaging with the IAMs on a conceptual level. Marten (2011), who compares IAMs with a simple upwelling diffusion energy balance model, comes closest to discussing the physical assumptions embedded in these IAMs. When we make the physics of IAMs the focus of our inquiries, however, we find that there are challenges both in interpreting many of the physical parameters and equations and in relating them between models—a problem that cannot be solved by computational experiments alone. This acts as a barrier to informed dialogue between the economists and policy makers who use these IAMs, and the scientists who study the physical processes of climate change and estimate the physical parameters embedded within them. In turn, this hampers our ability to distill today's best understanding into clear policy-relevant information (Revesz et al. 2014), and there is now an active debate on how to make these models more transparent (Nature Climate Change Editorial 2015; Rosen 2015; Smith et al. 2015; National Academies of Sciences, Engineering, and Medicine 2016).

We focus here on understanding the differences in model structure at a conceptual level, which provides us with clear physical interpretations of differences in model inputs and outputs. We show how this approach can be used as a basis for successively eliminating physically based differences in model output, and we illustrate our approach by comparing the temperature forecasting components of DICE, PAGE, and FUND. These IAMs use different physical models to forecast temperatures, different numerical approaches to solving these models, different values of physical parameters, and different terminologies to identify physical parameters. To systematize the comparison, we begin by relating their temperature forecasting components to simple energy balance models of the climate system. This provides a means to sort out terminological differences and a way to attribute forecast differences to specific modeling choices. Our results highlight several explicit and implicit physical assumptions in these IAMs that deserve greater scrutiny and input from the scientific community. These underscrutinized assumptions give rise to considerably different visions of the human and economic costs of climate change.

Our aim is not to show which IAM is best, but rather to provide the research community with the tools to identify physical reasons why they disagree. Our results can be used as a basis for running physically identical baseline simulations in all three IAMs, so as to separate the economic sources of disagreement from the physical. Our hope is that this

---

[2] For further examples, see Warren et al. (2006), Hof et al. (2012), Ackerman and Munitz (2012), and Gillingham et al. (2015).

additional transparency enables a more informed cross-disciplinary debate about which differences reflect genuine disagreements, thereby helping to focus attention and resources where they can most improve climate policy.

THE PHYSICS OF IAMS. A natural starting point for any effort to increase the transparency of DICE, FUND, and PAGE is to look at how they forecast annual-mean, global-mean surface temperature anomalies T. One reason is that T is the main input for computing the economic damages of climate change in these IAMs (see online supplement). A second reason is that there is a relatively strong physical basis for calculating T (see "The physics of temperature forecasts" sidebar). It can be shown that the temperature forecasting equations in all three IAMs derive from one of two simple physical climate models. PAGE and FUND are based on a single-equation climate model (Andrews and Allen 2008; Senior and Mitchell 2000; Dickinson 1986), while DICE is based on a two-equation model that divides Earth's thermal mass into one box for the deep oceans and another for the upper oceans, land surface, and atmosphere (Held et al. 2010; Geoffroy et al. 2013; see sidebar on "Forecasting temperature with IAMs" for more information). The link to these underlying physical models provides a means to translate IAM parameters into quantities studied in the scientific literature as well as to parameters in other IAMs. The translation key is presented in Table 1.

It is noteworthy that all three IAMs include a parameter for climate sensitivity (CS)—the equilibrium surface warming that results from a doubling of atmospheric CO_2 concentration—despite this parameter not appearing explicitly in the simple physical models from which they are derived (see "The physics of temperature forecasts" sidebar). In the physical science literature, energy balance models of this type are almost always written in terms of a "feedback parameter" (Dickinson 1982, 1986; Randall et al. 2007; Flato et al. 2013), often represented by λ as in the sidebar on "The physics of temperature forecasts." This parameter captures the combined consequences of temperature-dependent physical processes—such as changing surface albedo, clouds, increasing atmospheric water vapor, changing lapse rate, direct radiative effects, and so on—for the radiative balance at the surface. In the IAMs and in the economics literature, it is common to use the equilibrium climate sensitivity for this purpose. In equilibrium, the climate sensitivity is inversely proportional to the feedback parameter, but the trajectory to equilibrium inevitably consists of time-varying feedbacks (Senior and Mitchell 2000; Gregory et al. 2015)—a particularly obvious example is the consequence of sea ice decline on surface albedo whose feedback becomes zero once all the sea ice is gone. There are therefore many time series of λ, and correspondingly many temperature trajectories, that would lead to the same equilibrium temperature but that can produce different conclusions regarding the economic value of additional mitigation. It is of course possible to use the equilibrium relationship to back out the single feedback parameter value that relates to a given value of equilibrium climate sensitivity and then apply this in every period. This is what the IAMs do.

The IAMs are likely set up this way because of the focus on the equilibrium climate sensitivity in the physical science literature (Flato et al. 2013; Collins et al. 2013), and indeed, the values of CS in these IAMs closely reflect scientific estimates.

THE PHYSICS OF TEMPERATURE FORECASTS

One of the simplest climate change models considers the climate system as a box for which the rate of change in energy content is equal to changes in incoming radiation balanced by changes in outgoing radiation. This is often represented by Eq. (SB1) (Andrews and Allen 2008; Senior and Mitchell 2000; Dickinson 1986), in which the change in incoming radiation is taken to be the radiative forcing F (owing principally to changes in atmospheric greenhouse gas concentrations), and the change in outgoing radiation is taken to be proportional to the global-mean, annual-mean surface temperature anomaly T with a constant or proportionality λ known as the feedback parameter; C_{eff} is the effective heat capacity of the system:

$$C_{eff} \frac{dT}{dt} = F - \lambda T . \qquad \text{(SB1)}$$

The largest contribution to the effective heat capacity in Eq. (SB1) is the heat reservoir of the upper oceans. A simple extension of this model is to allow for the diffusion of heat to the deep oceans by adding a second box, the "deep ocean," which exchanges heat with the surface (or upper ocean) according to a one-dimensional heat transfer equation (Held et al. 2010; Geoffroy et al. 2013). This gives us the system

$$C_{up} \frac{dT}{dt} = F - \lambda T - \beta \left(T - T^{LO} \right)$$
$$C_{deep} \frac{dT^{LO}}{dt} = \beta \left(T - T^{LO} \right) . \qquad \text{(SB2)}$$

T^{LO} is the change in temperature for the deep ocean. The full parameter definitions are listed below Table 1.

FORECASTING TEMPERATURE WITH IAMS

IAMs forecast annual-mean, global-mean surface temperature anomaly T as a function of radiative forcing F. DICE uses the following pair of equations to calculate the temperature in each period t:

$$T_t = T_{t-1} + \xi_1 \left[F_t - \xi_2 T_{t-1} - \xi_3 \left(T_{t-1} - T_{t-1}^{LO} \right) \right]$$

$$T_t^{LO} = T_{t-1}^{LO} + \xi_4 \left(T_{t-1} - T_{t-1}^{LO} \right),$$

where T^{LO} is the change in lower ocean temperature. FUND forecasts T with the following equation:

$$T_t = \left(1 - \frac{1}{\phi}\right) T_{t-1} + \frac{1}{\phi} \frac{CS}{5.35 \ln(2)} F_t.$$

PAGE calculates T as follows:

$$T_t = T_{t-1} + \left[\frac{SENS}{FSLOPE_1 \ln(2)} F_t - T_{t-1} \right] \left(1 - e^{-(Y_t - Y_{t-1})/FRT} \right).$$

The full parameter definitions, as provided in each model's documentation, are listed below Table 1.

These temperature anomaly forecasting equations appear quite different at first glance, but each can be derived from a simple physical model of the climate system. One can show that the FUND and PAGE equations correspond to the one-box climate model in the sidebar on "The physics of temperature forecasts," while DICE corresponds to the two-box model (see online supplement for derivation). This also allows us to reexpress the many variously defined and described IAM parameters in terms of underlying physical quantities (see Table 1).

The most likely value of the CS in these IAMs are clustered around the IPCC's Fourth Assessment Report's modal estimate of 3°C (IPCC 2007; FUND, 3°C; PAGE, 2.54°C; DICE-2010, 3.2°C; DICE-2013, 2.9°C). FUND and PAGE, which are both designed to be run many thousands of times with different parameter values, effectively draw values of the equilibrium climate sensitivity from right-skewed distributions similar to those collected in Bindoff et al. (2013). In FUND there is a 74% chance that the CS lies in the IPCC's likely range (≥66% chance) of 1.5°–4.5°C (Bindoff et al. 2013); in PAGE it is a 93% chance.

Although the IAMs do a reasonable job capturing uncertainty about the equilibrium climate sensitivity, a physical understanding of these equations suggests that the material question is uncertainty in the feedback parameter. The difference is important for at least two reasons. First, the feedback parameter might be expected to vary with time or with CO_2 concentrations (Senior and Mitchell 2000; Meinshausen et al. 2011a), which means that uncertainty about the equilibrium climate sensitivity need not well reflect our uncertainty about temperature changes in coming decades and centuries.[3] Second, the transformation between the equilibrium climate sensitivity and the feedback parameter should account for uncertainty in the radiative forcing that results from CO_2 doubling ($F_{2\times CO_2}$). The IAMs assume very similar values for $F_{2\times CO_2}$ to those commonly cited in the scientific literature [$F_{2\times CO_2} \approx 3.7$ W m^{-2} (Myhre et al. 1998); FUND, 3.71; PAGE, 3.81; DICE-2010 and DICE-2013, 3.8], but they take no account of uncertainty in this quantity and therefore potentially understate uncertainty about the feedback parameter. The effective $F_{2\times CO_2}$ across phase 5 of the Coupled Model Intercomparison Project (CMIP5) ensemble is presented by the IPCC as having a 90% uncertainty of ±0.8 (Flato et al. 2013).

To date, sensitivity analyses of IAMs have focused on uncertainty in the equilibrium climate sensitivity (Ackerman et al. 2010; Dietz 2011; Pycroft et al. 2011; Gillingham et al. 2015). A physical perspective suggests that future development and sensitivity analyses of these IAMs might benefit from engaging with wider research on uncertainties in feedbacks, including studies on individual physical feedback processes (Boucher et al. 2013) and studies that quantify the time and state dependency of feedbacks in GCMs (Senior and Mitchell 2000; Meinshausen et al. 2011a).

Table 1 also reveals a noteworthy difference between the IAMs. FUND and PAGE express the thermal inertia of the system in terms of an *e*-folding time (third line of the table), while DICE uses heat capacities (although the parameters are not presented this way, see fifth and sixth lines of the table and online supplement for details). A system's heat capacity is the amount of energy input needed to heat it by 1°C, while its *e*-folding time, like a half-life, tells us how long it takes for the system to pass a certain distance (measured by the number *e*) on its way to a new equilibrium temperature.[4] Both concepts

---

[3] Confusingly, some studies that quantify the time and state dependency of feedbacks in GCMs reinterpret their results in terms of a time-dependent "effective climate sensitivity," which is distinct from the equilibrium climate sensitivity used in IAMs (Meinshausen et al. 2011b; Senior and Mitchell 2000; Murphy and Mitchell 1995).

[4] As applied to climate, the effective heat capacity is the amount of energy input per square meter of surface area that leads to an increase in global mean surface temperature of 1°C.

TABLE 1. IAM parameter equivalence. Cells are left empty if the parameter in question is not well defined within an IAM. We have also indicated where the parameter in question has been hard-coded in the IAM. The descriptions and values of IAM parameters below are those used in the IAM's documentation. Units are listed as specified in documentation and omitted whenever unspecified. The parameter values below are modal estimates whenever the documentation specifies a probability density function.

Physical/numerical	DICE	FUND	PAGE
CS	η	CS	SENS
$F_{2\times CO_2}$	$F_{2\times CO_2}$	$5.35\ln(2)$	$FSLOPE_i \ln(2)$
e-folding time	—	$\phi\Delta t$	$FRT \times 3.15576 \times 10^7$
C_{eff}	—	$\dfrac{5.35\ln(2)}{CS}\phi\Delta t$	$\dfrac{FSLOPE_i \ln(2)}{SENS}FRT \times 3.15576 \times 10^7$
C_{up}	$\dfrac{\Delta t}{\xi_1}$	—	—
C_{deep}	$\dfrac{\xi_3\Delta t}{\xi_4}$	—	—
λ	ξ_2	$\dfrac{5.35\ln(2)}{CS}$	$\dfrac{FSLOPE_i \ln(2)}{SENS}$
β	ξ_3	—	—
Δt	(Hard-coded)	(Hard-coded)	$(Y_t - Y_t - 1 \times 3.15576 \times 10^7$

Physical parameters:
CS: Climate sensitivity (K)
$F_{2\times CO_2}$: Radiative forcing for a doubling of CO_2 (W m^{-2})
e-folding time: e-folding time (s)
C_{eff}: Effective heat capacity of the climate system (J m^{-2} K^{-1})
C_{up}: Effective heat capacity of the surface (J m^{-2} K^{-1})
C_{deep}: Effective heat capacity of deep ocean (J m^{-2} K^{-1})
λ: Feedback parameter (W m^{-2} K^{-1})
β: Heat transfer coefficient between the upper and lower ocean (W m^{-2} K^{-1})
Δt: Length of time step t (s)

DICE parameters (DICE-2010/DICE-2013R):
$\xi_1 = 0.208/\xi_1 = 0.098$: Speed of adjustment parameter for atmospheric temperature
ξ_2: The ratio of increased forcing from CO_2 doubling ($F_{2\times CO_2}$ = 3.8/$F_{2\times CO_2}$ = 3.8) to the climate sensitivity ($\eta = 3.2/\eta = 2.9$)
$\xi_3 = 0.31/\xi_3 = 0.088$: Coefficient of heat loss from atmosphere to oceans
$\xi_4 = 0.05/\xi_4 = 0.025$: Coefficient of heat gain by deep oceans
Δt: Hard-coded time step of 10 yr/5 yr (3.15576 × 10^8 s/1.57788 × 10^8 s)

FUND parameters:
CS ~ $\Gamma(k = 6.48, \theta = 0.54)$: Climate sensitivity (°C)
$\phi = \max(-42.7 + 29.1 \times CS + 0.001 \times CS^2, 1)$: e-folding time (yr)
Δt: Hard-coded time step of 1 yr (3.15576 × 10^7 s)

PAGE parameters:
$FSLOPE_i$ = 5.5: Slope of radiative forcing equation
FRT ~ $Triang(a =10, b =65, c = 30)$: Feedback response time, also referred to as the half-life of global warming (yr)
TCR ~ $Triang(a =1, b =2.8, c = 1.3)$: Transient climate response

$$\text{SENS} = \dfrac{TCR}{\left[1 - \dfrac{FRT}{70}\left(1 - e^{-70/FRT}\right)\right]}: \text{Climate sensitivity (°C)}$$

Y_t: Calendar year at the start of period t

measure the climate's thermal inertia, but the difference in terminology and representation complicate like-with-like comparison between the models, while arguably having more profound consequences as well. Heat capacities relate more closely to well-understood physical properties of matter (mainly water in the oceans) than do e-folding times. Even though their representative values on a planetary scale are uncertain and time dependent (principally as a consequence of ocean circulation), heat capacities are easier to interpret in terms of fundamental energy flows and can also be linked more directly to physical observations.

How does this play out in the actual values used by the models? In FUND, the modal value of the e-folding time of global warming is 44 yr ($\phi\Delta t$ = 44 yr), while PAGE assumes the most likely value is 30 yr (FRT = 30 yr). DICE uses a single value for each parameter rather than an ensemble, and using the expressions

in Table 1 we can deduce that DICE-2010 and DICE-2013 implicitly assume e-folding times of 43.3 and 41.5 yr, respectively, in their initial simulation periods (see online supplement). Expressed as effective heat capacities, these IAMs assume that it takes 1.72 GJ (FUND), 1.42 GJ (PAGE), 1.62 GJ (DICE-2010), and 1.73 GJ (DICE-2013) per square meter of surface area, respectively, to raise the average surface temperature by 1°C. All of these estimates are at the high end in comparison to Frame et al. (2005), which provide a central estimate of about 0.8 GJ m^{-2} K^{-1}, with a maximum 5%–95% confidence interval of 0.1–2.05 GJ m^{-2} K^{-1}. The effective heat capacity implied by the distribution sampled by FUND exceeds the central Frame et al. (2005) estimate with a probability of 0.85 and exceeds the top end of the 90% confidence region with a probability of 0.53. In PAGE, the corresponding probabilities are 0.96 and 0.08. Populating the lower end of the range would tend to produce a greater spread of temperature forecasts in IAMs, since lower heat capacities result in more rapid warming (Calel et al. 2015).

There is a further point to note here. The climate sensitivity and e-folding time are not chosen independently in the ensemble models FUND and PAGE. FUND draws a value of the CS from a Gamma distribution and computes the e-folding time as a quadratic function of the CS. PAGE, on the other hand, draws values of the e-folding time and the transient climate response from two triangular distributions and then computes the CS. In both cases, the CS and e-folding time become positively correlated, although not to the same degree (FUND, 0.99; PAGE, 0.58). A correlation between these quantities is appropriate—it arises from the underlying physical model [Eq. (SB1)]— but the functions used in FUND and PAGE go further and impose a correlation between the climate sensitivity and the effective heat capacity (FUND, 0.9; PAGE, –0.3). The processes underlying these quantities are largely physically unrelated so one might question whether such a relationship is desirable. Estimates of the climate sensitivity and the effective heat capacity in the scientific literature can be correlated (Frame et al. 2005; Andrews and Allen 2008), yet it is worth highlighting that this is a statistical correlation that results from trying to fit two parameters to a single estimate of past warming. To the extent that the underlying physical processes may change in the future such relationships need not reflect the values, we might wish to sample to explore future uncertainty. In any case, the equations describing the relationship between the equilibrium climate sensitivity and e-folding time in FUND and PAGE represent additional physical assumptions over and above the temperature forecasting equations.

The different representations of the climate system also mean that, while FUND and PAGE have effective heat capacities that are constant in time, DICE does not. DICE has constant heat capacities for each of its two constituent boxes (one box representing the upper ocean, land surface, and atmosphere and a second representing the deep ocean) and consequently has a time-varying effective heat capacity for the surface temperature. Under the restriction that the model parameters are constant, the IAMs will therefore behave differently even if they begin with values that represent identical physical behavior at the surface.

Reexpressing the IAM parameters in physical terms facilitates comparisons, particularly with GCMs (Geoffroy et al. 2013; Flato et al. 2013), and reveals that the differences between these IAMs do not typically reflect the uncertainties found in the scientific literature. Thus, whether the range of climate forecasts from these IAMs (or distributions from individual IAMs) were bigger or smaller than that found in CMIP5, one might reasonably be concerned that they do not represent the current state of knowledge. Modeling choices underpinning how the feedback parameter is calculated tend to understate scientific uncertainty, while heat capacities seem to be systematically overestimated and their uncertainty underestimated.

Reexpressing IAM parameters in terms typically found in the physical science literature also highlights that some of the IAMs' parameters depend on physical quantities that few papers evaluate directly. Representative heat capacities and heat transfer coefficients (as used in DICE), for instance, could potentially be calculated quite easily from observational datasets by those who collect and study them (e.g., Domingues et al. 2008; Levitus et al. 2012; Durack et al. 2014), if only they were known to be of value for these kinds of policy assessments. Making the IAMs more transparent reveals the policy value of carrying out what might often be relatively simple calculations with the latest datasets in hand.

By relating the temperature forecasting components of these IAMs to simple physical climate models, we are able to identify discrepant physical assumptions more easily. Even if the different assumptions were just as plausible, understanding that they are a source of differing conclusions is valuable, not least in pointing us toward relevant scientific literature to inform model development and in directing future research efforts. Next, we show how this understanding also allows us to quantify the effects of physically based disagreements between IAMs.

QUANTIFYING THE DIFFERENCES. We have seen that DICE, FUND, and PAGE make different assumptions both about how to represent the climate

system and about the values of underlying physical parameters. Our preliminary comparisons with the scientific literature do not give us confidence that the range of assumptions within or across these IAMs corresponds to scientific uncertainty. To narrow this gap, one might start by substituting the IAMs' assumptions with ones from the literature. This can and has been done for a few of the more straightforward issues we have raised, such as uncertainty about the value of the equilibrium climate sensitivity (Marten 2011), but in this particular case the distributions in FUND and PAGE seem reasonable to start with. Many of the more subtle physical issues we have discussed cannot be satisfactorily addressed in this way, unfortunately, because of the present gulf between the IAMs and the information available in the scientific literature.

A more productive line of inquiry in the short term is to compare the IAMs to each other, as they are. This will return a conservative answer to the question of whether physical differences between these IAMs are large enough that they could make a difference in economic assessments of climate change. In addition, this comparison is an opportunity to put our translation key to use. Using the simplest underlying physical model as a common reference point allows us to selectively eliminate differences in parameter values or model structure, which provides a general method for pinpointing and quantifying the consequences of differences in specific physical assumptions. This method can be used in future studies to run the IAMs with standardized physical assumptions, which would allow them to separate physical and economic uncertainties and to conduct baseline runs that are more easily comparable across studies.

To isolate differences in the temperature forecasting component in these models, we solve their temperature equations using identical forcing time series. This produces the same temperature forecasts as running the full IAMs but constraining them to have identical economies and carbon cycles. Figure 1 shows the IAMs' temperatures forecast up to the year 2300 under four representative concentration pathways (RCPs) used in CMIP5 (Meinshausen et al. 2011b).[5] Even

with identical forcing assumptions, the IAMs produce substantively different temperature forecasts. Under the lowest-forcing scenario, the temperature anomaly differs by 0.29°C between the highest and lowest forecast by the year 2100, rising to 0.45°C by 2150, and thereafter the models begin to converge slightly. Under higher-forcing scenarios, the forecasts tend to diverge more slowly, but ultimately more dramatically. The temperature anomaly differs by up to 0.39°C by the year 2100 (0.22°C under RCP4.5, 0.17°C under RCP6, and 0.39°C under RCP8.5). By 2200 the difference exceeds 0.50°C (0.55°C under RCP4.5, 0.69°C under RCP6, and 0.77°C under RCP8.5), and by 2300 it reaches a maximum of 1.61°C (under RCP8.5). As one would expect, these ranges are somewhat smaller than the 5%–95% uncertainty ranges reported for the CMIP5 ensemble (as they only reflect uncertainty in modal parameter values; see sidebar on "Comparison with CMIP5" for more information), but the more relevant question is whether these differences are large enough that they could be significant for the sort of economic questions that these models were constructed to answer.

In the RCP8.5 scenario, with Nordhaus's damage function (Nordhaus and Sztorc 2013) (which is considered conservative), and with a discount rate of $r = 5\%$ (considered high), the net present value (NPV) of future damages associated with the most damaging forecast is $6 trillion (U.S. dollars) more than that associated with the least damaging forecast.[6] For comparison, the average NPV of damages across the four IAMs is roughly $54 trillion under the same assumptions. So in this case the spread is roughly 10% of the mean. This is as low as it goes in any scenario we look at; in other scenarios the spread rises to more than 50% of mean damages. The difference between IAM temperature forecasts can clearly cause

---

[5] While DICE produces a single forecast, FUND and PAGE incorporate uncertainty about parameter values and therefore produce distributions of forecasts. For expositional purposes, our discussion focuses on forecasts based on modal parameter values. This simplification tends to be conservative, and we discuss alternative specifications in the online supplement. Accounting for uncertainty in the parameter values tends to increase the spread across models, though this is left for the online supplement since it makes model comparison considerably more complicated and difficult to interpret.

[6] In a simple optimal growth model, the discount rate can be written as the sum of the pure rate of time preference ρ and the product of economic growth g and the elasticity of marginal utility η: $r = \eta g + \rho$. Conventionally, η is set somewhere in the range from 1 to 1.5, which tends to produce a discount rate of circa 5% when ρ is around 2%–3%. In this section, we obtain the same discount rate by setting $\rho = 5\%$ and letting $\eta = 0$. The assumption that $\eta = 0$ means that each dollar is worth as much as the next—a linear utility function. In the absence of a linear utility function, damages become a function of additional assumptions about population growth and economic inequality, which substantially complicates computation and interpretation. For the sake of completeness, the analysis in this section is repeated with a more common nonlinear utility function in the online supplement. The conclusions remain the same.
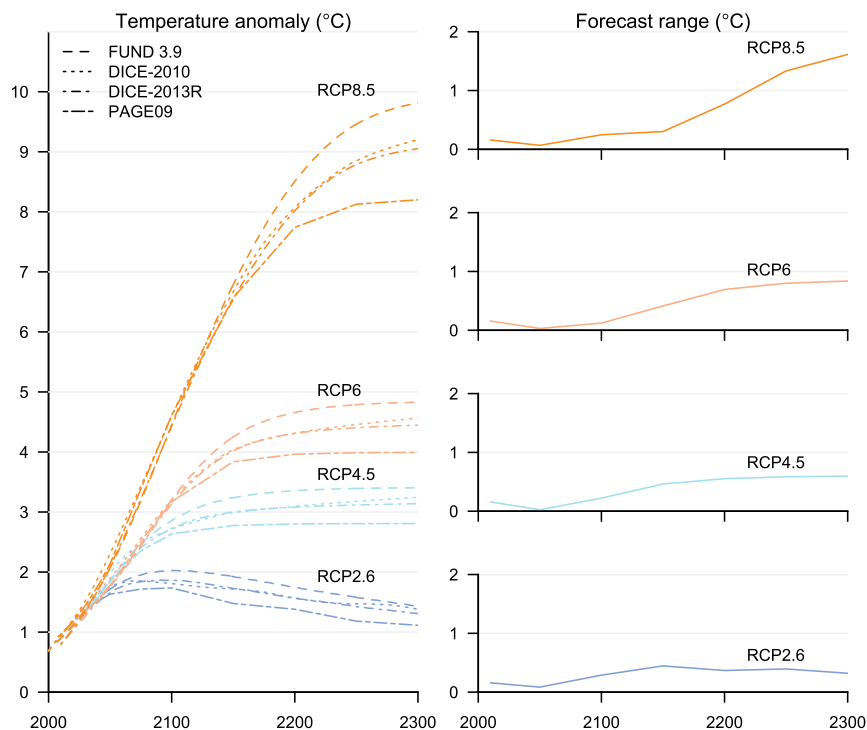
FIG. 1. Temperatures under RCP forcing scenarios: (left) The forecasted temperature anomaly for the choice of parameter values deemed most likely in each IAM. DICE specifies a single value for each parameter, while FUND and PAGE specify probability distributions for some parameters. For comparison purposes, we take the modal value of relevant parameters in FUND and PAGE. We have also included two versions of DICE, owing to the extensive use of older versions of DICE in economic assessments of climate change. The temperature forecasting components of DICE-2010 and DICE-2013 are structurally similar, but use different parameter values. (right) The maximum difference between forecasts over time.

the average NPV of future net output. Table 2 presents equivalent values for the different scenarios, different discount rates, and two frequently discussed damage functions. For instance, in the same scenario as above, but with Weitzman's more cautionary damage function (Weitzman 2012), the spread in damages is nearly $30 trillion, or 1.14% of average NPV of net output.

Table 2 leads us to make two observations. First, the differences in temperature forecasts across IAMs imply differences of several trillion dollars even under conservative assumptions. In proportional terms, these differences can amount to over 5% of future output with Norhdaus's damage function and nearly 30% with Weitzman's damage function. The range of the results highlight the importance of disagreements about discount rates and damage functions, as the economics literature has

significant variation in the magnitude of predicted damages, but a still more relevant comparison is perhaps economic output. Current annual world output is circa $80 trillion. Assuming that output would grow by 2% yr^{-1} absent climate damages, the stream of future gross output is worth $2,800 trillion in today's dollars, or just under $2,750 trillion if we subtract average damages.[7] The spread of damage forecasts, $6 trillion, then corresponds to 0.22% of

done already. But it also shows that, when operating within the domain of plausible economic parameter values, the physically based disagreements may have large enough economic consequences to warrant much greater attention. This attention is further justified by the fact that many physically based disagreements and uncertainties are qualitatively different to disagreements about economic parameters, such as the discount rate. They reflect the current level of understanding of the physical world (which we might hope to improve on in the future) rather than, for instance, ethical judgements about the relative social value of present and future output. Even if the physically based disagreements were of smaller consequence, there may still be great value in focusing more attention on the physics of IAMs because these disagreements may be easier to resolve. Doing so makes it easier to disentangle the qualitatively different economic uncertainties that can drive IAMs to distinct climate policy recommendations.

---

[7] In principle, we would want to measure the discounted value of an infinite stream of future output, which is a convergent sum as long as the discount rate is sufficiently high. Discounting provides a systematic way to compare the value of present and future output. Output in the very distant future will carry very little weight in the economic calculations. In practice, we truncate the economic calculations in 2500 since this is as far into the future as the RCPs extend. The precise year at which the calculations are truncated matters less and less the further into the future it occurs.

It is tempting, but difficult, to compare the range of IAM outputs with the range of forecasts made by the GCMs in the CMIP5 ensemble. The 5%–95% model ranges for the period 2081–2100 in that ensemble—1.4°C (RCP2.6), 1.5°C (RCP4.5), 1.7°C (RCP6), and 2.2°C (RCP8.5) (IPCC 2013, Table SPM.2)—are greater than the differences we present between the IAMs, but the ranges measure fundamentally different things. The IAM ranges we report pertain to forecasts that use modal parameter values; the modes of distributions that are chosen, presumably, to be representative of the best scientific understanding of these parameters. CMIP5, by contrast, is an ensemble of available models—an ensemble of opportunity (Allen and Stainforth 2002). It reflects uncertainty in the parameters themselves, which is of course greater than the uncertainty in the modal parameter values, so one would expect the range of CMIP5 forecasts to be

significantly greater than the range of modal forecasts from the IAMs. At the same time, the CMIP5 ensemble is not a random sample thought to be representative of scientific uncertainty, as summarized by the distributions of parameters such as climate sensitivity (Allen et al. 2006); in particular it does not reflect the tails of such a distribution. Therefore, to the extent that FUND and PAGE (the two IAMs that explore parameter uncertainty)

choose parameter distributions to reflect the best scientific understanding, one would expect the 5%–95% range of individual IAM distributions to be wider still than the CMIP5 5%–95% range. Figure SB1 demonstrates these general features, although it is worth noting that the CMIP5 ensemble occasionally encompasses a wider range than the individual IAM distributions, which suggests that these IAMs may be too constrained.

► Fig. SB1. Comparison with CMIP5: The gray lines show the CMIP5 ensemble of forecasts, downloaded from the KNMI Climate Explorer (https://climexp.knmi.nl/, 28 Sep 2016). The black lines show the same modal IAM forecasts as in Fig. 1, rebased to the global mean temperature in FUND over the period 1986–2005 to provide greater comparability with the way the CMIP5 ensemble is presented by the IPCC (IPCC 2013). The shaded areas show the 5%–95% confidence regions obtained from ensemble runs of FUND and PAGE when we draw a large random sample from the parameter distributions specified by each (see online supplement for details).
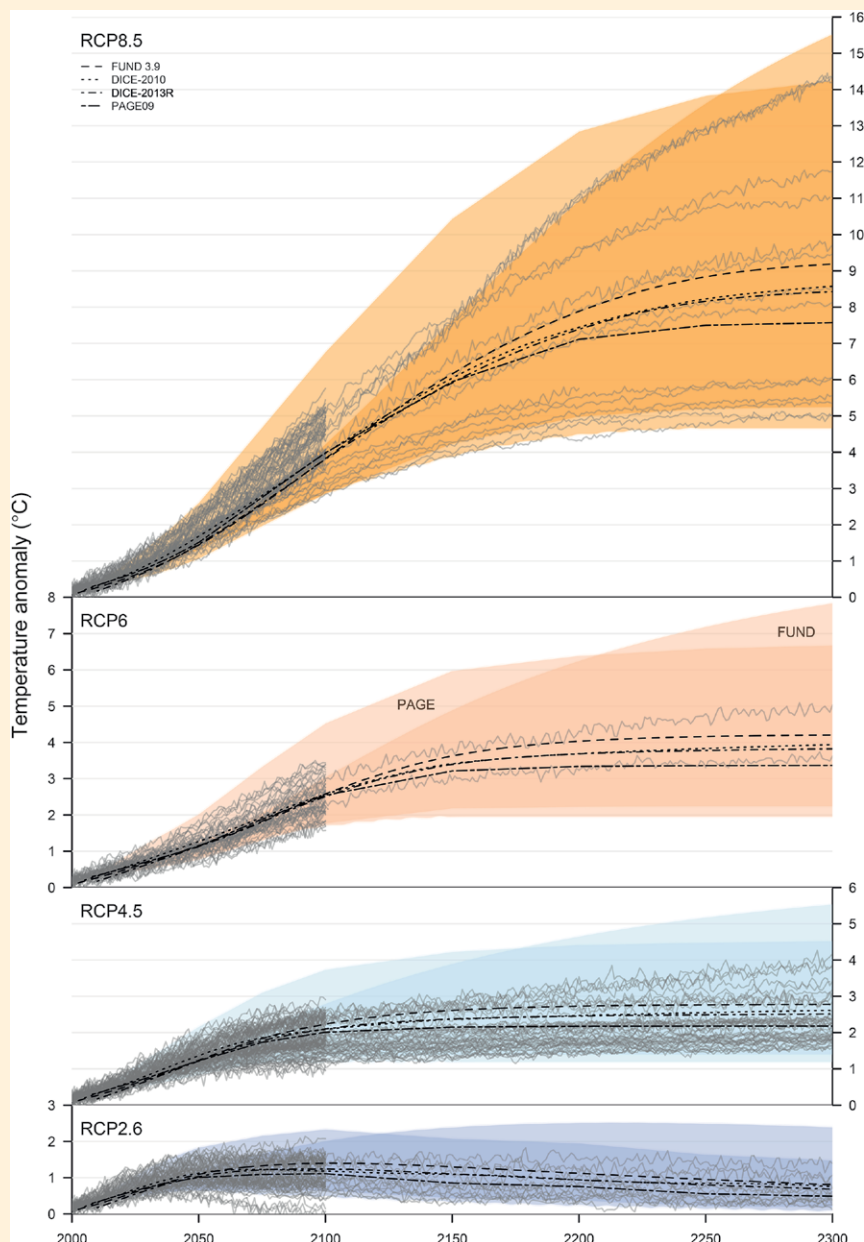
TABLE 2. Differences in economic damages as a percentage of the NPV of future output: In each period, a fraction $D(T)$ of gross economic output is lost as a result of climate change, where $D(\cdot)$ is taken to be either Nordhaus's or Weitzman's damage function. Assuming that economic output would grow at a rate of g yr^{-1}, absent climate damages, and that we have a constant pure rate of time preference ρ, we can write the NPV of the flow of damages as $\mathbb{D} = \sum_{s=0}^{s}[1/(1+\rho)^s]Y_0 D[T(s)](1+g)^s$, where s indexes time and Y_0 is current world output. Note that this calculation effectively uses a linear utility function, so that ρ is equal to the discount rate r. The table gives the largest difference across models, $\max(\mathbb{D}_{DICE-2010}, \mathbb{D}_{DICE-2013R}, \mathbb{D}_{PAGE}, \mathbb{D}_{FUND}) - \min(\mathbb{D}_{DICE-2010}, \mathbb{D}_{DICE-2013R}, \mathbb{D}_{PAGE}, \mathbb{D}_{FUND})$, as a percentage of the average NPV of future output less damages. The numbers in this table assume $g = 2\%$. See online supplement for additional discussion and scenarios.

$r =$	2%	3%	5%	2%	3%	5%
		RCP2.6			RCP4.5	
Nordhaus	0.19	0.17	0.09	0.81	0.41	0.11
Weitzman	0.18	0.17	0.09	1.73	0.80	0.17
		RCP6			RCP8.5	
Nordhaus	1.54	0.64	0.13	6.59	1.91	0.22
Weitzman	10.20	3.35	0.39	28.65	7.18	1.14

Second, the economic consequences of temperature forecast differences are determined by a complicated interaction of discounting and damage functions. The discount rate determines when forecast differences are important, while the curvature of the damage function determines at what temperature differences matter most. Usually we think of high discount rates as putting less weight on the future costs of climate change and lower discount rates as implying higher damage costs. However, in a low-forcing scenario most of the damages occur relatively early on, so a lower discount rate can imply smaller proportional damages. This effect is somewhat muted in Table 2, but we see a clear example of it in the online supplement where discounting at 2% results in smaller proportional damages than discounting at 3% (RCP2.6, $g = 3\%$). More generally, larger differences in temperature forecasts are not always more economically significant. For certain combinations of discount rates and damage functions, the damages associated with a smaller absolute difference in temperature forecasts can be amplified relative to future output if this difference occurs at a particular time and at a temperature where the damage function is strongly curved. This provides a further reason to be cautious about dismissing even apparently small forecast differences.

Even when the discrepancies between IAMs appear relatively tolerable in aggregated and discounted terms, though, it is important to remember that these differences nevertheless represent substantially different visions of a warming world. For instance, as the temperature rises we may pass several feared climatic tipping points (Lenton et al. 2008). The horizontal distances in Fig. 1 give us a range of estimates for when we may pass these thresholds (Joshi et al. 2011).

Under RCP6, for instance, we may cross the tipping point taken to be indicative of triggering a dieback of the Amazon rain forest (+3°–4°C) as much as 20–117 yr sooner or later, depending on whether we look at FUND or DICE, while the PAGE forecast allows for the possibility that we may never cross the tipping point if it is at +4°C. The range is even greater when we consider parameter uncertainty within FUND and PAGE (see online supplement). The range across IAMs of when a tipping point is reached is generally wider in lower-forcing scenarios, if the IAMs reach the relevant threshold at all. Table 3 summarizes the differences between IAMs in terms of date ranges when different global mean temperature anomaly thresholds are crossed and the associated climatic nonlinearities that may be triggered [using estimates from Lenton et al. (2008)].

Increasing temperatures have also been associated with an increased frequency and ferocity of hurricanes (Emanuel 2005), droughts (Kelley et al. 2015), depletion of freshwater resources (Jiménez Cisneros et al. 2014), and so forth, and recent studies have found highly nonlinear responses to rising temperatures in everything from economic productivity (Burke et al. 2015b) to agricultural yields (Schlenker and Roberts 2009; Welch et al. 2010), to human migration (Bohra-Mishra et al. 2014), and violent conflict (Hidalgo et al. 2010; Burke et al. 2015a). We cannot easily translate differences in global mean temperature into additional deaths from conflict, or some such metric, but the differences between IAMs represent futures where we may have several more or fewer decades before these consequences are upon us, decades that could be spent forestalling or better adapting to such developments. Now that IAMs are
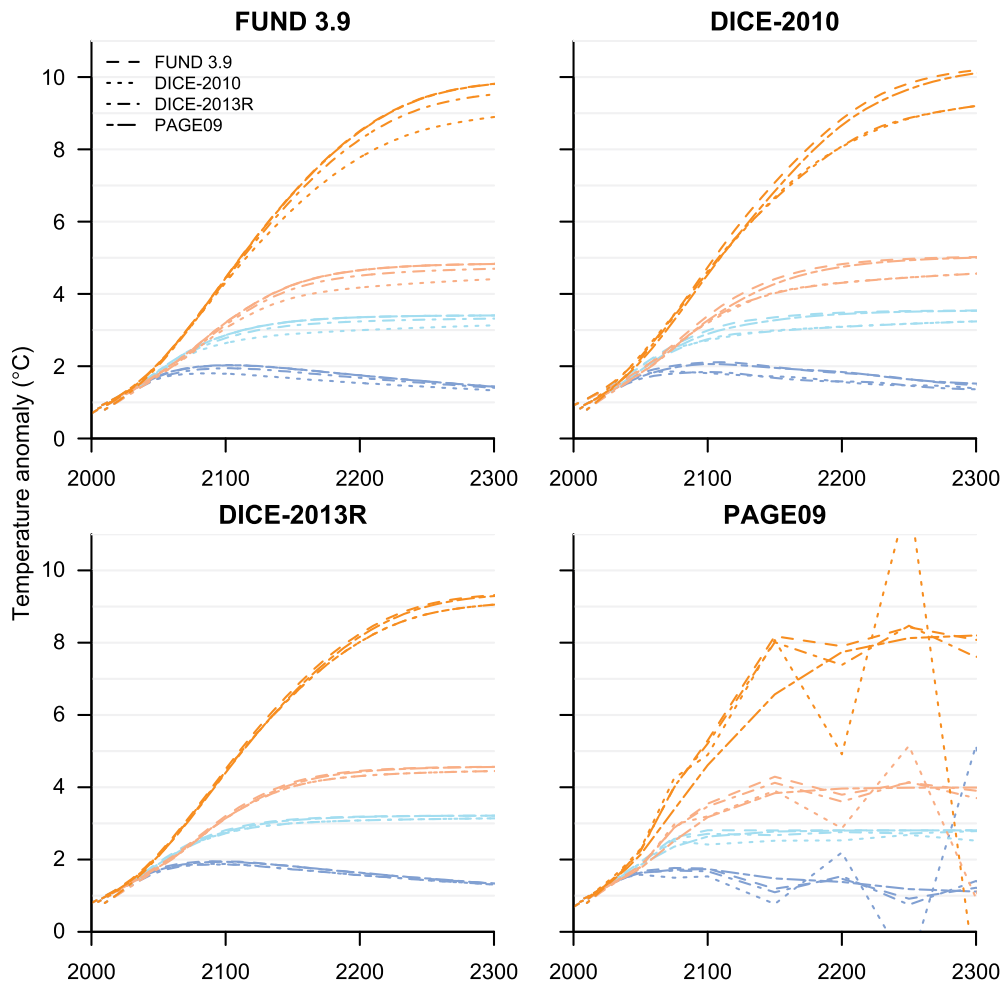
FIG. 2. Temperatures with comparable parameter values: As in the left panel of Fig. 1, but within each subplot all models are run with the same physical parameter values at the start of their simulation. The parameter values used are those taken as default in the model indicated in the title of each subplot.

beginning to be extended to incorporate these other aspects of a warming world, such as climatic tipping points (Lontzek et al. 2015), the differences in the physics of IAMs may come to matter more and more.

In sum, whether we measure temperatures (Fig. 1), damages (Table 2), or time before tipping points are passed (Table 3), there remain substantial discrepancies between IAM forecasts even after eliminating differences in radiative forcing. The main source of these discrepancies is the different assumptions made about the values of the physical parameters.

RESOLVING THE DIFFERENCES. Using our physical translation key (Table 1 and online supplement), each model's temperature components can be run with parameter values that reflect the physical assumptions implicit in the other models. Figure 2 replicates the left panel of Fig. 1 once for the parameter values chosen in each model.

Three things are worth noting about Fig. 2. First, the temperature forecasts are generally more in line with each other when the IAMs use the same initial values for physical parameters. The FUND and PAGE temperature forecasts now coincide, except for tiny differences due to initialization. DICE-2010 and DICE-2013R start off almost identical to FUND and PAGE, but notable differences begin to appear after a few decades. In the first three panels of Fig. 2, the highest and lowest predictions now differ by at most 0.3°C in 2100. In some instances the differences still exceed 0.5°C by 2200, and by 2300 the IAMs can differ by nearly 1°C in the highest-forcing scenario. Differences in parameter values sometimes compensate for structural model differences, so eliminating the variation in the initial parameter values while holding model structure fixed occasionally increases the spread of forecasts. This approach disentangles choices of parameters and structure and, thus, provides a way

TABLE 3. The timing of tipping points: Columns 2–5 list the ranges of dates across IAMs at which global-mean, annual-mean temperature anomaly crosses the thresholds indicated in column 1 (1°, 2°, 3°, 4°, 5°, and 6°C). Column 6 lists a selection of macroclimatic events along with the estimated ranges of global mean temperature, which is taken to be indicative of triggering those events, taken from Lenton et al. (2008, their Table 1). See Lenton et al. (2008) for a full description of the events and the methodology for estimating the ranges. The events are grouped and sorted by the associated ranges, so they only roughly line up with the temperature thresholds in column 1. Column 7 summarizes the key climatic impacts thought to be associated with these events, also taken from Lenton et al. (2008). This table defines the temperature anomaly relative to 1980–99 to provide greater comparability with how Lenton et al. (2008) defines temperature thresholds.

Tempera-ture (°C)	Timing range across IAMs				Tipping element (global warming)	Key impacts
	RCP2.6	RCP4.5	RCP6	RCP8.5		
1	2040–45	2038–41	2040–44	2036–38	Arctic summer ice (+0.5°–2°C)	Amplified warming
2	—	2079–97	2078–84	2062–64	Greenland ice sheet (+1°–2°C)	Sea level +2–7 m
3	—	—	2113–33	2083–86	Amazon rain forest (+3°–4°C)	Biodiversity loss
4	—	—	2179–2296[a]	2102–09	West Antarctic ice sheet (+3°–5°C)	Sea level +5 m
					Atlantic thermohaline circulation (+3°–5°C)	Regional cooling, sea level, ITCZ shift
5	—	—	—	2122–34	Increase of Sahara/Sahel and West African monsoon (+3°–5°C)	Increased carrying capacity
					Boreal forest dieback (+3°–5°C)	Biome switch
6	—	—	—	2145–64	Increased amplitude of El Niño–Southern Oscillation (+3°–6°C)	Drought in SE Asia and elsewhere

[a] Only FUND and DICE-2010 cross the 4°C threshold in scenario RCP6, at least before the year 2500, so that the date range in this table understates the range of uncertainty captured by this set of forecasts.

to quantify the effect of any specific disagreement about the value of a physical parameter, so long as the underlying model structures are sufficiently similar to allow translation of parameters. In this case one can also eliminate all differences in initial parameter values simultaneously, which raises the prospect for IAM studies to include baseline assessments with standardized physical parameter values.

Second, although the models are now initialized with identical physical parameter values and are subjected to identical forcings, DICE produces systematically lower temperature forecasts over time.[8] The slower response of DICE stands in contrast to some earlier IAM comparisons (e.g., van Vuuren et al. 2011) where instead PAGE and FUND were identified

as responding comparatively slowly. The slow initial response of FUND described by van Vuuren et al. (2011) is likely a consequence of parametric assumptions (earlier versions of FUND had longer e-folding times and therefore a slower response), while our finding highlights the relative importance of physical parameter values and model structure. DICE, unlike FUND and PAGE, includes a lower ocean into which surface heat can escape. Figure 2 shows that for the same initial physical parameter values all models respond in a similar fashion at first but that the difference in model structure, in terms of ocean heat uptake, becomes increasingly important over time. In physical terms, the structural difference causes the effective heat capacity in DICE to vary over time,

[8] The more rapid response of DICE-2013 compared with DICE-2010 (see the first three subplots of Fig. 2) is due to a 70% reduction in the heat loss to the deep ocean. This comes about by a 70% reduction in the heat transfer coefficient between the upper and lower ocean accompanied by a 70% reduction in the heat capacity of the deep ocean. The change in heat capacity is a consequence of a change in the "coefficient of heat loss from the atmosphere to oceans," which is not compensated for by a change in the "coefficient of heat gain by the deep ocean." These changes imply substantial changes in the physical world—one might conceptualize it as substantial decreases in the strength of the global ocean conveyor accompanied by the removal of millions of cubic kilometers of deep ocean water. The change appears simply to be a consequence of "calibration," but no physical explanation is given for it as far as the authors have been able to determine. Equally puzzling changes from previous versions of other IAMs can also be found (e.g., the recent reduction in e-folding time from 66 to 44 yr in FUND). Changes of this magnitude go beyond fine tuning and this is a good example of where climate scientists could offer useful input and where that community might want to be involved in the discussion of suitable values.
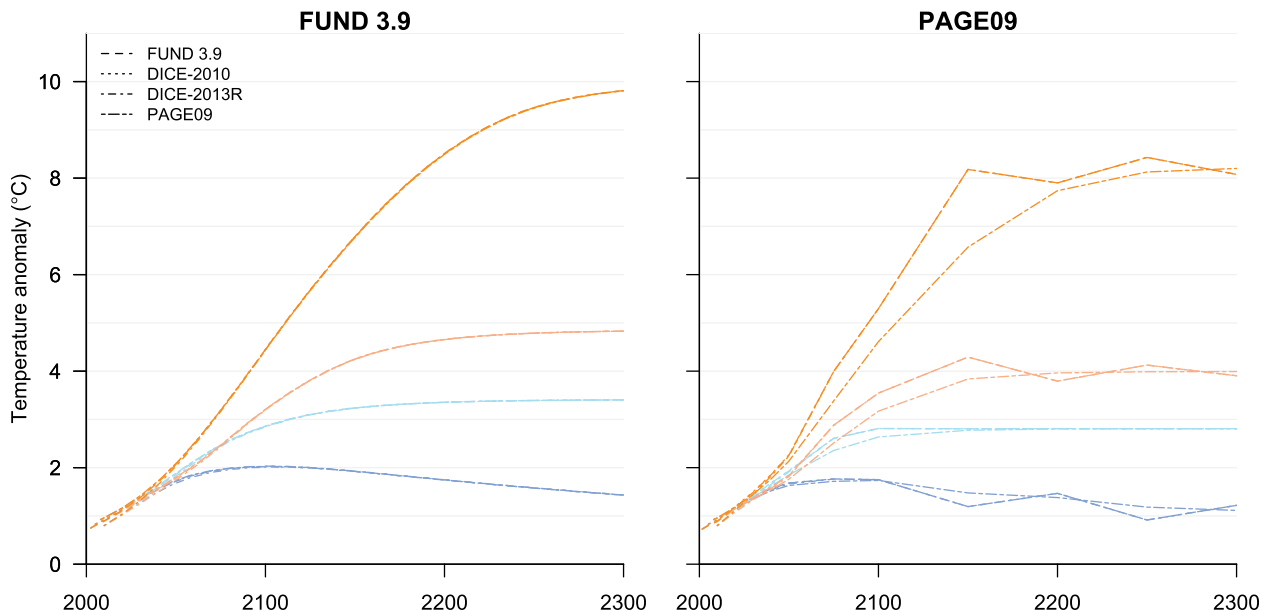
FIG. 3. Temperatures with constant effective heat capacities: As in the leftmost and rightmost panels of Fig. 2, but varying the IAM parameters in DICE-2010 and DICE-2013R so that the implied effective heat capacity is constant over time. See online supplement for a full explanation of how the IAM parameters are calculated.

while it is constant in FUND and PAGE (see online supplement). This suggests a method for quantifying the consequences of structural differences between the IAMs—provide time-varying effective heat capacities in FUND and PAGE to mimic those seen in a similar scenario in DICE, or let the parameters in DICE vary over time to produce a more stable effective heat capacity. In the limit, we can eliminate the structural difference by letting the parameters in DICE vary so that the implied effective heat capacity is constant. Figure 3 shows that this almost entirely removes the remaining differences between temperature forecasts. This approach can be used in future IAM studies to conduct baseline assessments with identical physical assumptions, both about the values of physical parameters and model structure.

Third, a striking aspect of Fig. 2 is the numerical instability of DICE and FUND when run with the long time steps used by PAGE. It is this same instability, coupled with slightly different dates of initialization, that prevents DICE and FUND from exactly matching PAGE in the second panel of Fig. 3. This is arguably less interesting since these models do not typically attempt to use such long time steps. Nevertheless, it does illustrate the potential for instability and therefore the value of maintaining the time step Δt as a separate parameter that can be varied to facilitate running model ensembles and sensitivity analyses where parameters can take values in the tail of their probability distribution (see online supplement discussion of numerical representation).

These IAMs represent three legitimately different sets of modeling choices, and policy makers can benefit from access to IAMs that reflect the range of economic and scientific uncertainty and disagreement. Linking the physics of these IAMs on a conceptual level allows us to see how well the range of IAMs corresponds to the range of scientific uncertainty and to have a much more precise and transparent discussion about the nature and consequences of those different modeling choices. Many scientific uncertainties remain—for instance, there is uncertainty about the consequences for climate change of present-day ocean circulation patterns and the current rate of ocean heat uptake (Hawkins et al. 2016), and how these might vary in the future (Gregory et al. 2015)—but uncertainty about these quantities is qualitatively different to uncertainties about many economic parameters, such as the rate of time preference. It is possible to bring the physics of these IAMs into the open and to standardize physical assumptions across IAMs in a rigorous way. This allows us to selectively eliminate differences in forcing, parameter values, and model structure (summarized in Fig. 4), which makes it easier for model calibrations to treat these model differences separately so that we do not conflate qualitatively different types of uncertainty.

GOING FORWARD TOGETHER. While it is useful for policy makers to have access to different IAMs, it is also critical that these models are openly

examined and reexamined by economists and the scientific community to ensure that they reflect our current best understanding, and so that we can identify areas of disagreement where research would be most valuable.

While economists have honed in on a handful of critical economic modeling choices in recent years, the broader scientific community has been less successful distilling and comparing the physics of the three IAMs that have probably been the most influential in policy debates—DICE, FUND, and PAGE. Our approach has been to start from a physical understanding of the climate component of these IAMs and then successively eliminate sources of disagreement—first standardizing assumptions about forcing, then the values of physical parameters, then model structure, leaving only numerical representation differences (further discussed in the online supplement). Our aim is not to say which IAM is better for what purpose but to provide a common framework that allows systematic comparison and informed debate about their differences.

We find that, even with identical economies, these three IAMs produce substantially different climate change forecasts. These differences may not be as significant as the choice of discount rate or damage function, but they nevertheless correspond to many trillions of dollars of damage and should not be casually dismissed. By tracing the climate components of IAMs back to basic physical models, we can clarify the sources of these disagreements, pointing to where

in the scientific literature one might look to resolve them. For instance, we find that there is significant disagreement about the effective heat capacity across IAMs, and the models even disagree over how to incorporate heat-capacity-like information. Our approach allows one to selectively remove variation arising from a specific source of disagreement and, thereby, to quantify its importance.

Our analysis also suggests a need for caution in future IAM development. Many consequential differences can be embedded in just one or two equations, even in equations with a strong and sometimes identical physical basis. If we race to make IAMs yet more complex, it is likely that it also becomes more difficult to understand why they disagree. In the earlier example of ocean dynamics, for instance, FUND and PAGE are built on a one-box model, while DICE uses a two-box model to capture the fact that there is heat loss to the deep ocean (and thus a time-varying effective heat capacity) as the surface warms. A diffusive ocean (Hansen et al. 1985) would capture this effect better, and a three-dimensional ocean circulation model better still. It is tempting to steadily increase model complexity, maybe to the point of embedding economic components into three-dimensional GCMs, or by separating the physical and economic components and using GCM output as direct inputs to economic evaluations (Burke et al. 2015b). Yet, it may be more beneficial to consider strategies to
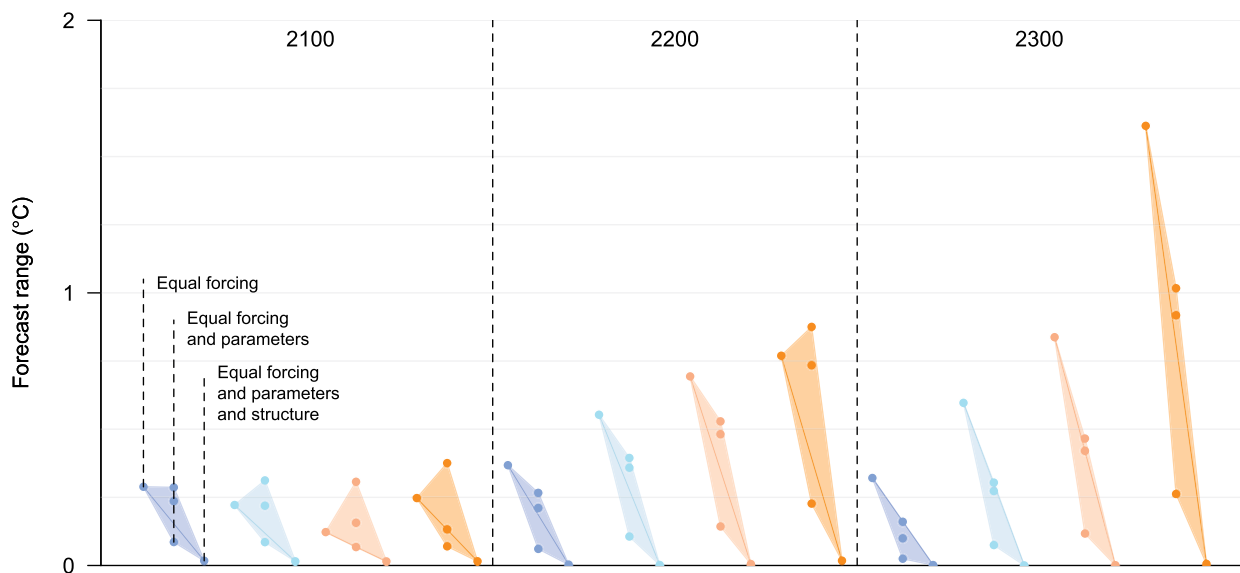

Fig. 4. Differences between IAM forecasts: Each polygon encompasses the forecast ranges (from Figs. 1, 2, and 3) at a particular point in time (2100, 2200, or 2300) associated with a given RCP scenario (2.6, 4.5, 6, or 8.5); individual ranges are represented by the colored dots. The forecast ranges from the rightmost panels of Figs. 2 and 3 are omitted owing to their numerical instability. The different colors denote different RCPs, as in Fig. 1. Going left to right within each polygon, we see that differences in IAM temperature forecasts diminish when we standardize forcing, parameter values, and model structure. The diagonals of the polygons are shown to help guide the eye.

improve the dynamic behavior of IAMs without increasing model complexity. An alternative to adding a second box (increasing model complexity) would be to represent changing ocean heat uptake by introducing a time-dependent or temperature-dependent heat capacity. This allows the one-box model to replicate the dynamics of the more complex two-box model or of GCMs and other models of plausible responses. Such an approach may be less intellectually satisfying because it makes the process exogenous to the model. On the other hand, it creates a requirement for those researchers to consult with ocean dynamicists who are arguably best placed to judge the range of plausible ocean circulation responses and their effect on energy flows, which might ultimately produce results that better reflect current understanding of the relevant processes; better perhaps than could be achieved with extra equations. A similar argument can be made for climate sensitivity and the feedback parameter, both of which are taken to be constant in these IAMs but would in practice be state dependent and, therefore, time dependent (Senior and Mitchell 2000; Williams et al. 2008). There are clearly advantages and disadvantages to any strategy—whether adding equations, time-varying parameters, or some other approach—and we urge only that further complexity be justified in the context of the aims of the model and after consideration of all options, not simply because there is greater computational capacity to implement it.

Our findings and method can be put to at least one immediate practical use. Having made explicit the link between these IAMs and the underlying physical models of the climate system, we have gained the ability to translate between IAMs using a common physical language. This translation key will enable multimodel policy assessments to run all three models with physically comparable baseline scenarios, which would isolate the economic sources of disagreement.

In the longer term, we believe that increasing the visibility of the link between the physical sciences and the economic analyses will also help the scientific community focus more keenly on those unresolved questions that loom largest in policy assessments. We also hope that making these IAMs more accessible to the scientific community will invite further scientific expertise into the IAM community, so that economic assessments of climate change reflect the latest physical understanding of the climate system.

REFERENCES

Ackerman, F., and C. Munitz, 2012: Climate damages in the FUND model: A disaggregated analysis. *Ecol. Econ.*, **77**, 219–224, doi:10.1016/j.ecolecon.2012.03.005.

——, E. A. Stanton, and R. Bueno, 2010: Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE. *Ecol. Econ.*, **69**, 1657–1665, doi:10.1016/j.ecolecon.2010.03.013.

Allen, M. R., and D. A. Stainforth, 2002: Towards objective probabalistic climate forecasting. *Nature*, **419**, 228–228, doi:10.1038/nature01092a.

——, D. Frame, J. A. Kettleborough, and D. A. Stainforth, 2006: Model error in weather and climate forecasting. *Predictability of Weather and Climate*, T. Palmer, Ed., Cambridge University Press, 391–427.

Andrews, D. G., and M. R. Allen, 2008: Diagnosis of climate models in terms of transient climate response and feedback response time. *Atmos. Sci. Lett.*, **9**, 7–12, doi:10.1002/asl.163.

Anthoff, D., and R. Tol, 2013: FUND—Climate Framework for Uncertainty, Negotiation and Distribution. [Available online at https://github.com/fund-model/fund.]

——, and ——, 2014: The Climate Framework for Uncertainty, Negotiation and Distribution (FUND), technical description, version 3.9. FUND Doc., 26 pp. [Available online at www.fund-model.org/versions.]

Bindoff, N., and Coauthors, 2013: Detection and attribution of climate change: From global to regional. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 867–952.

Bohra-Mishra, P., M. Oppenheimer, and S. M. Hsiang, 2014: Nonlinear permanent migration response to climatic variations but minimal response to disasters. *Proc. Natl. Acad. Sci. USA*, **111**, 9780–9785, doi:10.1073/pnas.1317166111.

Boucher, O., and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis*, T.

F. Stocker et al., Eds., Cambridge University Press, 571–657.

Burke, M., S. M. Hsiang, and E. Miguel, 2015a: Climate and conflict. *Annu. Rev. Econ.*, **7**, 577–617, doi:10.1146/annurev-economics-080614-115430.

——, ——, and ——, 2015b: Global non-linear effect of temperature on economic production. *Nature*, **527**, 235–239, doi:10.1038/nature15725.

Calel, R., D. A. Stainforth, and S. Dietz, 2015: Tall tales and fat tails: The science and economics of extreme warming. *Climatic Change*, **132**, 127–141, doi:10.1007/s10584-013-0911-4.

Clarke, L., and Coauthors, 2014: Assessing transformation pathways. *Climate Change 2014: Mitigation of Climate Change*, O. Edenhofer et al, Eds., Cambridge University Press, 413–510.

Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1029–1136.

Dasgupta, P., 2007: Commentary: The Stern Review's economics of climate change. *Natl. Inst. Econ. Rev.*, **199**, 4–7.

Dickinson, R. E., 1982: Modelling climate changes due to carbon dioxide increases. *Carbon Dioxide Review*, W. C. Clark, Ed., Clarendon Press, 101–133.

——, 1986: How will climate change? *The Greenhouse Effect, Climatic Change and Ecosystems*, B. Bolin et al., Eds., John Wiley and Sons, SCOPE Report, Vol. 29, 206–270.

Dietz, S., 2011: High impact, low probability? An empirical analysis of risk in the economics of climate change. *Climatic Change*, **108**, 519–541, doi:10.1007/s10584-010-9993-4.

Domingues, C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker, and J. R. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature*, **453**, 1090–1093, doi:10.1038/nature07080.

Durack, P. J., P. J. Gleckler, F. W. Landerer, and K. E. Taylor, 2014: Quantifying underestimates of long-term upper-ocean warming. *Nat. Climate Change*, **4**, 999–1005, doi:10.1038/nclimate2389.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688, doi:10.1038/nature03906.

EPA, 2014: Regulatory impact analysis for the proposed carbon pollution guidelines for existing power plants and emission standards for modified and reconstructed power plants. U.S. Environmental Protection Agency Tech. Rep. EPA-452/R-14-002, 376 pp. [Available online at www.epa.gov/sites /production/files/2014-06/documents/20140602ria -clean-power-plan.pdf.]

Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

Frame, D. J., B. B. B. Booth, J. A. Kettleborough, D. A. Stainforth, J. M. Gregory, M. Collins, and M. R. Allen, 2005: Constraining climate forecasts: The role of prior assumptions. *Geophys. Res. Lett.*, **32**, L09702, doi:10.1029/2004GL022241.

Geoffroy, O., D. Saint-Martin, D. J. L. Olivié, A. Voldoire, G. Bellon, and S. Tytéca, 2013: Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments. *J. Climate*, **26**, 1841–1857, doi:10.1175/JCLI-D-12-00195.1.

Gillingham, K., and Coauthors, 2015: Modeling uncertainty in climate change: A multi-model comparison. National Bureau of Economic Research Working Paper 21637, 66 pp., doi:10.3386/w21637.

Greenstone, M., E. Kopits, and A. Wolverton, 2013: Developing a social cost of carbon for US regulatory analysis: A methodology and interpretation. *Rev. Environ. Econ. Policy*, **7**, 23–46, doi:10.1093/reep/res015.

Gregory, J. M., T. Andrews, and P. Good, 2015: The inconstancy of the transient climate response parameter under increasing CO_2. *Philos. Trans. Roy. Soc. London*, **A373**, 2054, doi:10.1098/rsta.2014.0417.

Hahn, R. W., and R. A. Ritz, 2015: Does the social cost of carbon matter? Evidence from US policy. *J. Leg. Stud.*, **44**, 229–248, doi:10.1086/680990.

Hansen, J., G. Russell, A. Lacis, I. Fung, D. Rind, and P. Stone, 1985: Climate response times: Dependence on climate sensitivity and ocean mixing. *Science*, **229**, 857–859, doi:10.1126/science.229.4716.857.

Hawkins, E., R. Smith, J. Gregory, and D. Stainforth, 2016: Irreducible uncertainty in near-term regional climate projections. *Climate Dyn.*, **46**, 3807–3819, doi:10.1007/s00382-015-2806-8.

Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Climate*, **23**, 2418–2427, doi:10.1175/2009JCLI3466.1.

Hidalgo, F. D., S. Naidu, S. Nichter, and N. Richardson, 2010: Economic determinants of land invasions. *Rev. Econ. Stat.*, **92**, 505–523, doi:10.1162/REST_a_00007.

Hof, A. F., C. W. Hope, J. Lowe, M. D. Mastrandrea, M. Meinshausen, and D. P. van Vuuren, 2012: The benefits of climate change mitigation in integrated assessment models: The role of the carbon cycle and

climate component. *Climatic Change*, **113**, 897–917, doi:10.1007/s10584-011-0363-7.

Hope, C., 2006: The marginal impact of CO_2 from PAGE2002: An integrated assessment model incorporating the IPCC's five reasons for concern. *Integr. Assess. J.*, **6**, 19–56.

——, 2011: The PAGE09 integrated assessment model: A technical description. Cambridge Judge Business School Working Paper, 44 pp. [Available online at www.jbs.cam.ac.uk/fileadmin/user_upload/research/workingpapers/wp1104.pdf.]

Interagency Working Group, 2010: Technical support document: Social cost of carbon for regulatory impact analysis—Under Executive Order 12866. United States Government Interagency Working Group on Social Cost of Carbon Tech. Rep., 51 pp.

——, 2015: Technical support document: Technical update of the social cost of carbon for regulatory impact analysis—Under Executive Order 12866. United States Government Interagency Working Group on Social Cost of Carbon Tech. Rep., 35 pp.

IPCC, 2007: Climate change 2007. IPCC Synthesis Rep., 52 pp. [Available online at www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.]

——, 2013: Summary for policymakers. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 3–29.

Jiménez Cisneros, B. E., T. Oki, N. W. Arnell, G. Benito, J. G. Cogley, P. Döll, T. Jiang, and S. S. Mwakalila, 2014: Freshwater resources. *Climate Change 2014: Impacts, Adaptation, and Vulnerability, Part A: Global and Sectoral Aspects*, C. B. Field et al., Eds., Cambridge University Press, 229–269.

Joshi, M., E. Hawkins, R. Sutton, J. Lowe, and D. Frame, 2011: Projections of when temperature change will exceed 2°C above pre-industrial levels. *Nat. Climate Change*, **1**, 407–412, doi:10.1038/nclimate1261.

Kelley, C. P., S. Mohtadi, M. A. Cane, R. Seager, and Y. Kushnir, 2015: Climate change in the Fertile Crescent and implications of the recent Syrian drought. *Proc. Natl. Acad. Sci. USA*, **112**, 3241–3246, doi:10.1073/pnas.1421533112.

Lenton, T. M., H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, 2008: Tipping elements in the Earth's climate system. *Proc. Natl. Acad. Sci. USA*, **105**, 1786–1793, doi:10.1073/pnas.0705414105.

Levitus, S., and Coauthors, 2012: World Ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, L10603, doi:10.1029/2012GL051106.

Lontzek, T. S., Y. Cai, K. L. Judd, and T. M. Lenton, 2015: Stochastic integrated assessment of climate tipping

points indicates the need for strict climate policy. *Nat. Climate Change*, **5**, 441–444, doi:10.1038/nclimate2570.

Marten, A. L., 2011: Transient temperature response modeling in IAMs: The effects of over simplification on the SCC. *Economics*, **5** (2011-18), 1–42, doi:10.5018/economics-ejournal.ja.2011-18.

Meinshausen, M., S. C. B. Raper, and T. M. L. Wigley, 2011a: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6—Part 1: Model description and calibration. *Atmos. Chem. Phys.*, **11**, 1417–1456, doi:10.5194/acp-11-1417-2011.

——, and Coauthors, 2011b: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241, doi:10.1007/s10584-011-0156-z.

Millner, A., and G. Heal, 2014: Resolving intertemporal conflicts: Economics vs politics. National Bureau of Economic Research Working Paper 20705, 44 pp., doi:10.3386/w20705.

Murphy, J. M., and J. F. B. Mitchell, 1995: Transient response of the Hadley Centre Coupled Ocean–Atmosphere Model to increasing carbon dioxide. Part II: Spatial and temporal structure of response. *J. Climate*, **8**, 57–80, doi:10.1175/1520-0442(1995)008<0057:TROTHC>2.0.CO;2.

Myhre, G., E. J. Highwood, K. P. Shine, and F. Stordal, 1998: New estimates of radiative forcing due to well mixed greenhouse gases. *Geophys. Res. Lett.*, **25**, 2715–2718, doi:10.1029/98GL01908.

National Academies of Sciences, Engineering, and Medicine, 2016: Assessment of approaches to updating the social cost of carbon: Phase 1 report on a near-term update. Committee on Assessing Approaches to Updating the Social Cost of Carbon and Board on Environmental Change and Society and Division of Behavioral and Social Sciences and Education and National Academies of Sciences, Engineering, and Medicine Tech. Rep., 72 pp. [Available online at www.nap.edu/catalog/21898/assessment-of-approaches-to-updating-the-social-cost-of-carbon.]

Nature Climate Change Editorial, 2015: IAM helpful or not? *Nat. Climate Change*, **5**, 81, doi:10.1038/nclimate2526.

Nordhaus, W. D., 2007: A review of the "Stern Review on the Economics of Climate Change." *J. Econ. Lit.*, **45**, 686–702, doi:10.1257/jel.45.3.686.

——, 2013: DICE-2013R model as of November 15, 2013. [Available online at www.econ.yale.edu/~nordhaus/homepage/Web-DICE-2013-April.htm.]

——, and P. Sztorc, 2013: DICE 2013R: Introduction and user's manual. 2nd ed. 102 pp. [Available online at www.econ.yale.edu/~nordhaus/homepage/documents/DICE_Manual_100413r1.pdf.]

Pycroft, J., L. Vergano, C. Hope, D. Paci, and J. C. Ciscar, 2011: A tale of tails: Uncertainty and the social cost of carbon dioxide. *Economics*, **5** (2011-22), 1–29, doi:10.5018/economics-ejournal.ja.2011-22.

Randall, D., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 589–662.

Revesz, R. L., P. H. Howard, K. Arrow, L. H. Goulder, R. E. Kopp, M. A. Livermore, M. Oppenheimer, and T. Sterner, 2014: Global warming: Improve economic models of climate change. *Nature*, **508**, 173–175, doi:10.1038/508173a.

Rose, S., D. Turner, G. Blanford, J. Bistline, F. de la Chesnaye, and T. Wilson, 2014: Understanding the social cost of carbon: A technical assessment. Electric Power Research Institute, 184 pp. [Available online at www.epri.com/#/pages/product/000000003002004657/.]

Rosen, R. A., 2015: IAMs and peer review. *Nat. Climate Change*, **5**, 390–390, doi:10.1038/nclimate2582.

Schlenker, W., and M. J. Roberts, 2009: Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc. Natl. Acad. Sci. USA*, **106**, 15 594–15 598, doi:10.1073/pnas.0906865106.

Senior, C. A., and J. F. B. Mitchell, 2000: The time-dependence of climate sensitivity. *Geophys. Res. Lett.*, **27**, 2685–2688, doi:10.1029/2000GL011373.

Smith, S. J., and Coauthors, 2015: Long history of IAM comparisons. *Nat. Climate Change*, **5**, 391–391, doi:10.1038/nclimate2576.

Stern, N., 2007: *The Economics of Climate Change: The Stern Review*. Cambridge University Press, 712 pp.

van Vuuren, D. P., and Coauthors, 2011: How well do integrated assessment models simulate climate change? *Climatic Change*, **104**, 255–285, doi:10.1007/s10584-009-9764-2.

Warren, R., C. Hope, M. Mastrandrea, R. S. J. Tol, N. Adger, and I. Lorenzoni, 2006: Spotlighting the impacts functions in integrated assessments. Research Rep. prepared for the Stern Review on the Economics of Climate Change, Tyndall Centre Working Paper 91.

Watkiss, P., and C. Hope, 2011: Using the social cost of carbon in regulatory deliberations. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 886–901, doi:10.1002/wcc.140.

Weitzman, M. L., 2007: A review of the "Stern Review on the Economics of Climate Change." *J. Econ. Lit.*, **45**, 703–724, doi:10.1257/jel.45.3.703.

——, 2012: GHG targets as insurance against catastrophic climate damages. *J. Public Econ. Theory*, **14**, 221–244, doi:10.1111/j.1467-9779.2011.01539.x.

Welch, J. R., J. R. Vincent, M. Auffhammer, P. F. Moya, A. Dobermann, and D. Dawe, 2010: Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures. *Proc. Natl. Acad. Sci. USA*, **107**, 14 562–14 567, doi:10.1073/pnas.1001222107.

Williams, K. D., W. J. Ingram, and J. M. Gregory, 2008: Time variation of effective climate sensitivity in GCMs. *J. Climate*, **21**, 5076–5090, doi:10.1175/2008JCLI2371.1.