Title
Characterizing and responding to uncertainty in climate change

Author
Lemoine, Derek Mark

Publication Date
2011

Characterizing and responding to uncertainty in climate change

by

Derek Mark Lemoine

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Energy and Resources

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Daniel M. Kammen, Chair
Professor Larry S. Karp
Professor Michael O'Hare
Assistant Professor Christian Traeger

Spring 2011

Characterizing and responding to uncertainty in climate change

Abstract

Characterizing and responding to uncertainty in climate change

by

Derek Mark Lemoine

Doctor of Philosophy in Energy and Resources

University of California, Berkeley

Professor Daniel M. Kammen, Chair

The development and analysis of climate policy proposals intertwine with the structure of knowledge and the possibility for changing it. Key questions concern the long-term interaction between policy, technology, infrastructure, and the earth system, but each of these components is deeply uncertain. This dissertation advances the description of knowledge about the climate system, the assessment of economic responses to climatic possibilities, and the development of policy that positions society to achieve long-term climate goals. It offers new paths to describing understanding of complex systems and to modeling optimal management under structural uncertainty.

The first chapter formalizes uncertainty about equilibrium climate change. Its hierarchical Bayes framework allows climate models to be incomplete and to share biases, and it shows how prior beliefs about models' completeness and independence interact with models' estimates of feedback strength to determine distributions for temperature change. When models might share biases, the results of additional models might tell us more about models' common structure than about the real-world processes they aim to represent. The most valuable information would then come not from related models but from alternate estimates that should carry a different set of unobservable biases. The possibility that models are wrong in common ways limits the degree to which models' estimates can narrow the probability distribution for feedback strength, which also limits our ability to rule out extreme climatic outcomes.

The second chapter empirically estimates a feedback that is especially difficult to model. Climate-carbon feedbacks (or carbon cycle feedbacks) describe the effect of temperature on carbon dioxide (CO_2). If they are positive, then not only does anthropogenic CO_2 cause warming via the greenhouse effect and earth system feedbacks, but this warming itself increases CO_2 and so causes further warming. Previous empirical work estimated a stronger feedback than did coupled climate-carbon cycle models. However, those empirical estimates were probably biased upwards while coupled models' estimates were primarily driven by a few ill-constrained parameters. This chapter attempts to obtain an unbiased estimate of

climate-carbon feedback strength by using variations in summer radiation in the Arctic (i.e., variations in orbital forcing) to identify the effect of temperature on CO_2 in 800 ky ice core records. It finds a range for climate-carbon feedbacks that is closer to coupled models' estimates than to previous empirical work. Since climate-carbon feedbacks are probably positive, temperature change projections tend to underestimate an emission path's consequences if they do not allow the carbon cycle to respond to changing temperatures.

The next three chapters assess economic responses to climate change in a policy-optimizing integrated assessment model, in games with long-lived investments into abatement capital, and in a cost-effectiveness model with multiple policy options stretching over long time horizons. The first of these chapters extends a well-known integrated assessment model to include the possibility of abrupt shifts in the climate system. It also changes the model's structure to make the decision-maker aware of uncertainty and of the possibility for learning over time, and it generalizes the welfare evaluation to reflect that uncertainty about temperature change is qualitatively unlike uncertainty about climate thresholds. It finds that tipping points can increase the near-term social cost of carbon by more than 50% when they raise climate sensitivity or make damages more convex. They have less of an effect when they increase the atmospheric lifetime of CO_2 or the quantity of non-CO_2 greenhouse gases. Allowing the policymaker to be differentially averse to consumption fluctuations over time and over risk increases the near-term social cost of carbon by 150%, with tipping point possibilities then increasing it by another 50%. The possibility of tipping points is more important for the social cost of carbon than is the ambiguity attitude the decision-maker uses in evaluating them.

The second of these climate economics chapters models the optimal emission tax when firms can adopt low-pollution technology that reduces abatement cost. The regulator anticipates this adoption but must set the tax before firms invest. In many cases, a linear emission tax cannot obtain both socially optimal investment and socially optimal emissions because the regulator either will set it inefficiently high to stimulate investment or will set it at an ex post optimal level that obtains inefficiently low investment. The difficulty is that an emission tax fixes both the incentive to invest and the incentive to abate, but these two goals rarely align perfectly when investment is lumpy. In contrast, tradable permits policies do not suffer this tension because the permit price responds automatically to realized investment. A numerical model then considers the ability of the regulator to select not only the level but also the duration of the tax. It shows that outcomes are still often socially inefficient. Further, the regulator will occasionally use a longer tax to obtain investment when firms expect their investments to lower the tax in the next period, but the cost of not being able to adjust the next period's tax limits the parameter space in which the longer tax is employed.

The fifth chapter constructs cost-effective dynamic policy portfolios of abatement, research and development (R&D), and negative emission technology deployment in order to achieve 21st century climate targets. It includes two types of stochastic technological change in a stylized numerical model and allows each type of technology to respond both to public R&D and to abatement policies. It compares worlds where negative emission technologies are and are not available, and it compares a world where the century's cumulative net emissions are constrained with a world in which threshold possibilities lead policy to constrain cumulative net emissions in each year during the century. It finds that R&D options are valuable and exercised but do not substitute for near-term abatement. The type of R&D undertaken depends on long-term emission goals because those determine the magnitude of future abatement. When the cumulative emission constraint is stringent, negative emission technologies substitute for near-term abatement and affect the type of R&D undertaken, but if threshold considerations eliminate the freedom to temporarily overshoot emission targets, negative emission technologies become less valuable. The availability of negative emission technologies provides a valuable option to partially undo previous emissions, but abatement also gains option value from increasing future flexibility to forgo reliance on negative emission technologies if the technology or climate prove problematic in the interim.

The concluding chapter directly connects uncertainty about climate change to uncertainty about the cost of achieving CO_2 targets. It shows how beliefs about technology, temperature, and damages interact to affect the cost-effectiveness of climate targets. It finds that the speed with which damages increase at higher temperatures is the most important of these factors. Both 450 parts per million (ppm) and 550 ppm CO_2 targets provide net benefits for quadratic damage functions that reduce annual output by less than the 1-2% estimated for 2.5°C of warming. Cubic damage functions support both CO_2 targets even if 2.5°C of warming only reduces output by 0.2% or less. More convex damage functions also reduce the importance of abatement cost uncertainty. significantly increase the range of damage functions that support these targets and decrease the importance of abatement cost uncertainty. In addition, because extreme feedback outcomes have little effect over the next decades, a thinner-tailed temperature distribution (resulting from optimistic prior beliefs about climate models' independence and biases) supports CO_2 targets under slightly less severe damages than does the thicker-tailed distribution (resulting from skepticism about climate models' independence and biases). Emission reductions hedge against greater societal sensitivity to temperature increases while exposing society to the upside of positive technology surprises.

The epistemology of complex systems in an out-of-sample world is a key motif. This dissertation advances knowledge of climate change and understanding of policy design in settings with limited ability to predict future changes or responses. Further work should seek a more unified framework for describing and acting on knowledge of evolving complex systems.

Family, friends, and places

Contents

List of Figures **iv**

List of Tables **vi**

1 Climate sensitivity distributions depend on the possibility that models share biases **1**

1.1 Introduction . . . 2
1.2 Previous approaches to developing distributions for temperature change . . . 2
1.3 Feedback analysis . . . 3
1.4 Methods: Hierarchical Bayes framework . . . 6
1.5 Data: Models' estimates of feedbacks . . . 11
1.6 Results: Posterior distributions . . . 16
1.7 Discussion . . . 22
1.7.1 Opportunities for learning . . . 22
1.7.2 Measures of temperature risk . . . 24
1.8 Conclusion . . . 26
1.9 Appendix: Estimating the carbon cycle feedback factor using coupled climate-carbon cycle models . . . 27

2 Paleoclimatic warming increased carbon dioxide concentrations **30**

2.1 Introduction . . . 30
2.2 Assessing feedback strength . . . 31
2.3 Methods: Estimated equations . . . 33
2.4 Results . . . 38
2.5 Discussion . . . 43
2.6 Appendix: Sources of bias in estimating climate-carbon feedbacks . . . 46
2.7 Appendix: Hierachical Bayes model for combining coupled models' output with empirical estimates . . . 48

3 Tipping points and ambiguity in the integrated assessment of climate change **52**
3.1 Introduction 53
3.2 Introducing tipping points and ambiguity aversion into DICE 57
3.3 Modeled tipping points 63
3.4 Results 65
3.5 Discussion 73
3.6 Appendix: Model calibration 76
3.7 Appendix: Model specification 81

4 Using the magnitude and duration of an emission tax to stimulate adoption of green technology **87**
4.1 Introduction 88
4.2 Analytic model 92
4.2.1 Description 92
4.2.2 The equilibria from which the regulator must select 94
4.2.3 The regulator's choice of equilibrium 97
4.2.4 The social optimality of equilibrium outcomes 104
4.2.5 Comparison to a quantity instrument 106
4.3 Multiperiod numeric model 107
4.3.1 Description and solution procedure 107
4.3.2 Numeric results 111
4.4 Discussion 114
4.5 Appendix: Additional proofs 117

5 The influence of negative emission technologies and technology policies on the optimal climate mitigation portfolio **125**
5.1 Introduction 126
5.2 Model of policy portfolio optimization 128
5.3 Results: Portfolio cost, robust actions, and critical parameters 133
5.4 Discussion: Policy implications 140
5.5 Appendix: Model parameterization 142

6 High temperature impacts are a dominant uncertainty in the evaluation of climate targets **147**
6.1 Three sources of uncertainty 148
6.2 Evaluating climate policy under uncertain technology, impacts, and science . 151
6.3 Results: Break-even output loss 154
6.4 The interaction of policy options with uncertainty 156

List of Figures

1.1 Temperature increases nonlinearly with feedback strength 5
1.2 Graph of the statistical model 8
1.3 Prior distributions 10
1.4 Feedback data 14
1.5 Posterior distributions for individual feedback factors 18
1.6 Posterior distributions for aggregate feedbacks and climate sensitivity 19
1.7 Posterior distributions as box plots 20
1.8 Posterior for the shared bias term 20
1.9 Posterior for climate sensitivity as compared to previous work 21
1.10 The potential for learning 23
1.11 Temperature risk 25

2.1 Temperature, CO_2, and forcing time series 36
2.2 Temperature and CO_2 relationships 37
2.3 Climate-carbon feedback results 40
2.4 Feedback strength over time 41
2.5 Posterior distributions 45
2.6 Prior distributions 50
2.7 Joint distribution of feedback strength and shared biases 51

3.1 DICE schematic 54
3.2 Threshold probability and hazard rate 61
3.3 Social cost of carbon and optimal CO_2 with an uncertain threshold 69
3.4 Abatement cost 70
3.5 Social cost of carbon with a certain threshold 70
3.6 Ambiguity aversion and the social cost of carbon 72
3.7 Transient feedback and CO_2 decay 78
3.8 Calibration 80
3.9 Optimal policy in a post-threshold world 86

4.1 Candidate equilibria 99

4.2 The regulator's choice of equilibrium 100
4.3 Optimal taxes as a function of investment probability 113
4.4 Investment probability as a function of incoming plant 114
4.5 Optimal taxes as a function of incoming plant 115

5.1 Relative cost of policy portfolios 134
5.2 Planned emission paths 136
5.3 Optimized probability of actions 138

6.1 Distributions for temperature in 2095 150
6.2 Schematic 152
6.3 Distribution of abatement costs from GCAM 152
6.4 Break-even output loss from 2.5°C of warming 155

List of Tables

1.1 Prior distributions 11
1.2 Feedback data 15
1.3 Percentile values for the aggregate feedback factor 26
1.4 Pairing coupled models to uncoupled models 29

2.1 Climate-carbon feedback results 39
2.2 Full results for the orbital forcing specifications 41
2.3 Results from univariate and non-instrumented regressions 42
2.4 Full results for decadal-scale feedbacks 44
2.5 Prior distributions 49

3.1 Tipping points modeled 55
3.2 Welfare specifications 63
3.3 Model runs 66
3.4 Optimized probability of crossing each threshold 71
3.5 Social cost of carbon in 2015 72
3.6 Parameterization of transition equations 83

4.1 Candidate equilibria 98
4.2 Parameterization 110

5.1 Decision variables and key parameters 129
5.2 Policy options 131
5.3 Parameter scenarios 146

Acknowledgments

ERG is a mysterious black box: it gathers a faint signal, adds dimensions of noise, and outputs a more robust signal that somehow points to two dimensions and all dimensions at once. It is a community with wide-ranging interests united by practical motivations. It is a training device for finding an interesting question, figuring out which field is the needed method, and following through to an answer. But most of all, ERG is an experiment in academic self-determinism. It opens doors around campus and permits new intellectual combinations, with the student able to decide whether and how far to access any of them. By providing intellectual space and opportunity, ERG's paths provide unlimited upside as well as the individual responsibility to realize it.

I owe special thanks to a few people. First, my family—David, Brenda, and Nathan—trusted in any unfamiliar path. They taught me to watch, listen, and question, and to treat mastery and knowledge as a function of effort and desire rather than of a credential's gloss. Next, my committee members each proved invaluable in their own way. Dan Kammen does not let details block vision and is unconstrained by past templates. Christian Traeger is a wonderful collaborator, a true friend, and generous with his time, code, and ideas. Larry Karp, one of the best people to think through any problem with, worked with me in an early leap of faith and later opened professional doors. Every conversation with Mike O'Hare bounces across his range of knowledge and interests: from garbage contracts in New Jersey to the true reason planes fly to negotiating for squirrels to, maybe, the topic ostensibly at hand. Two other members of the Berkeley community, Severin Borenstein and Max Auffhammer, always offered time they may not have had and need not have provided. Margaret Torn and John Harte gave unique training and careful feedback. Finally, Alex Farrell taught us all the value of the next challenge and the permeability of limits.

Several unique collections of people have made the last years special. ERG staff builds then greases every wheel, the students support each other in life, and the Dragons provide the highlight of every week. Many times I longed to slow time so as to savor the stage of life and appreciate the openness of the future. At the same time, the unfurling of paths powers its own pull forward. At least, by the time we get there, the next stage almost feels natural.

Chapter 1

Climate sensitivity distributions depend on the possibility that models share biases[1]

Uncertainty about biases common across models and about unknown and unmodeled feedbacks is important for the tails of temperature change distributions and thus for climate risk assessments. This paper develops a hierarchical Bayes framework that explicitly represents these and other sources of uncertainty. It then uses models' estimates of albedo, carbon cycle, cloud, and water vapor-lapse rate feedbacks to generate posterior probability distributions for feedback strength and equilibrium temperature change. The posterior distributions are especially sensitive to prior beliefs about models' shared structural biases: nonzero probability of shared bias moves some probability mass towards lower values for climate sensitivity even as it thickens the distribution's positive tail. Obtaining additional models of these feedbacks would not constrain the posterior distributions as much as would narrowing prior beliefs about shared biases or, potentially, obtaining feedback estimates having biases uncorrelated with those impacting climate models. Carbon dioxide concentrations may need to fall below current levels in order to maintain only a 10% chance of exceeding official 2°C limits on global average temperature change.

[1]Originally published as:
Lemoine, D.M. 2010. Climate sensitivity distributions depend on the possibility that models share biases. Journal of Climate 23(16): 4395-4415. doi:10.1175/2010JCLI3503.1

1.1 Introduction

The possibility of unexpectedly extreme climate change may be crucial for analyses that aim to operationalize policy targets or evaluate policy options. The tails of the temperature change distributions used in these analyses may therefore have special importance. For instance, policymakers aiming to avoid dangerous climate change now often focus on temperature targets such as avoiding 2°C of warming relative to pre-industrial levels (e.g., UNFCCC, 2009; MEF, 2009). Because of the uncertain connection between emission paths and temperature change, determining the implications for allowable emission paths requires defining the acceptable chance of missing the temperature targets. With benchmark risk metrics, allowable emission paths should have less than a 10% chance of overshooting the target, but such assessments require temperature change distributions that include the types of uncertainty important for tail probabilities.

The tails of temperature change distributions may also matter for economic assessments of greenhouse gas (GHG) policies because damages and utility are both nonlinear in temperature change (e.g., Weitzman, 2009). Willingness to pay to reduce the risks of climate change may therefore be sensitive to the positive tail of the climate sensitivity distribution (Newbold and Daigneault, 2009). Further, economic assessments may respond to the pervasive uncertainty in integrated assessment models of the economy and climate by forgoing the calculation of optimal emission paths in favor of determining the cost-effective actions needed to meet exogenous GHG constraints (Ackerman et al., 2009; Lemoine et al., 2011). These exogenous constraints may be determined by tolerance for climate change risks, which would again require temperature change distributions with well-characterized tails.

The positive tail of a temperature change distribution may be sensitive to types of uncertainty often excluded by previous work on temperature change probabilities. Much of the tail uncertainty is driven not just by uncertainty about the best climate model to use or about the best way to parameterize a given model but by uncertainty about features common across models, about our understanding of climate processes, and about how earth system processes in a future warming world may differ from those in the past periods for which we have data. After describing previous work and outlining the approach to feedback analysis, I propose and apply a hierarchical Bayes framework for developing posterior distributions for feedback factors that explicitly account for uncertainty about model completeness and shared structural biases. I then show how the implied temperature change distributions could inform risk assessments, policy targets, and future research into feedback processes.

1.2 Previous approaches to developing distributions for temperature change

A number of researchers have developed probability distributions for climate sensitivity, which gives the equilibrium temperature change produced by doubling carbon dioxide (CO_2)

concentrations from their pre-industrial level of approximately 280 ppm.[2] Most studies have reported a most likely value between 2 and 3.5°C, a 5% lower limit between 1 and 2°C, and an uncertain upper limit that often exceeds 6°C (Knutti and Hegerl, 2008). The current paper estimates posterior distributions for a parameter that is very similar to climate sensitivity but differs in allowing the carbon cycle to respond to the changing temperatures induced by exogenously doubled CO_2 concentrations.

Knutti and Hegerl (2008) reviewed the main approaches previous studies have taken towards constraining climate sensitivity: they have varied parameters in general circulation models and compared the results to climatic observations; they have estimated climate sensitivity from instrumental period data and from the longer, more variable time series available from paleoclimatic data; and the two studies of Annan and Hargreaves (2006) and Hegerl et al. (2006) have combined information from several constraints spanning both the instrumental and paleoclimatic records. Each of these approaches did not include several sources of uncertainty that may be important for tail risks because they did not explicitly include the possibility that models share biases or that past climates may be imperfect proxies for future climate change. In contrast to many previous studies (see Tebaldi and Knutti, 2007; Knutti et al., 2010), the proposed hierarchical Bayes methods recognize the possibility of structural biases shared across models, which limits the information gain from an unbounded increase in the number of models (compare Berliner and Kim, 2008). This statistical framework also includes uncertainty about climate models' completeness and about the similarity of the present and future higher-GHG world to the worlds represented by past climate observations.

1.3 Feedback analysis

The total temperature change from an increase in GHG concentrations depends not just on the direct effect of trapping additional outgoing infrared radiation but also on how the wider earth system responds to changing temperatures. For instance, temperature-induced changes in sea ice extent, vegetation, and water vapor content affect temperature by changing the earth's reflectivity and its ability to trap outgoing heat. Such changes are feedbacks that may amplify or diminish the effect of the direct radiative forcing (e.g., Roe, 2009). Because total temperature change can be highly sensitive to estimates of aggregate feedbacks, relatively small uncertainty about each feedback can translate into much greater uncertainty about total temperature change and into a significant possibility of extreme temperature change (e.g., Roe and Baker, 2007).

Roe (2009) described the framework for linear feedback analysis adopted here. As will be seen in the appendix, it is important to be clear about the system within which feedbacks operate, including the subsystem (known as the reference system or open system) to which

[2] The definition of climate sensitivity is ambiguous with regard to very fast feedbacks and to slow feedbacks (Knutti and Hegerl, 2008).

feedbacks return output as input (Stephens, 2005). Feedbacks are only meaningful in relation to a reference system that defines what happens in the absence of feedbacks. The reference system in the case of climate change is usually a blackbody planet that responds to a sustained increase in radiative forcing ΔR_f (such as from increased GHG concentrations) by adjusting its atmospheric radiation balance until it reaches a new equilibrium with a higher temperature and correspondingly greater outgoing radiation. The change in temperature for the blackbody planet in response to ΔR_f is $\Delta T_0 = \lambda_0 \Delta R_f$ for some λ_0 determined by the Stefan-Boltzmann law and atmospheric characteristics. This blackbody representation does not correspond to actual expectations of temperature change because the earth system includes processes that affect radiative forcing in the course of responding to temperature change. The temperature change occurring in the total earth system (or closed system) depends on how these feedbacks amplify or decrease the reference system's temperature change.

Feedbacks can be introduced in at least two equivalent ways (Roe, 2009). First, each of K feedback processes changes the radiative forcing by an amount proportional to the temperature change:

$$\Delta T = \lambda_0 \left(\Delta R_f + \sum_{k=1}^{K} c_k \Delta T \right) \tag{1.1}$$

In this representation, feedbacks are independent of each other, meaning that feedback i only affects feedback j via its effect on temperature change. Rearranging, we have:

$$\Delta T = \frac{\lambda_0 \Delta R_f}{1 - \sum_{k=1}^{K} \lambda_0 c_k} = \frac{\lambda_0 \Delta R_f}{1 - \sum_{k=1}^{K} f_k} = \frac{\lambda_0 \Delta R_f}{1 - F} \tag{1.2}$$

where the feedback factor for feedback k is $f_k = \lambda_0 c_k$ and the aggregate feedback factor F is the sum of the individual feedback factors: $F = \sum_{k=1}^{K} f_k$. The feedback factors affect temperature change nonlinearly (Figure 1.1). Second, we can use a Taylor series expansion to describe how temperature would change in response to changes in components of the earth system. This derivation underlies the estimation of feedback factors from climate models and from climatic records. Each feedback process k corresponds to a climate field α_k. This climate field changes in response to temperature, and changes in the climate field affect radiative forcing. Let ΔR_α be the change in radiative forcing due to the temperature-induced changes in the K climate fields. Then the Taylor series expansion yields:

$$\Delta R_\alpha = \frac{\mathrm{d}R_\alpha}{\mathrm{d}T} \Delta T + O(\Delta T^2) \approx \left\{ \sum_{k=1}^{K} \left[\left. \frac{\partial R}{\partial \alpha_k} \right)_{\alpha_{j, j \neq k}} \frac{\mathrm{d}\alpha_k}{\mathrm{d}T} \right] \right\} \Delta T \tag{1.3}$$

Equivalently to equation (1.1), we have:

$$\Delta T = \lambda_0 \left(\Delta R_f + \Delta R_\alpha \right) \approx \lambda_0 \left\{ \Delta R_f + \sum_{k=1}^{K} \left[\left. \frac{\partial R}{\partial \alpha_k} \right)_{\alpha_{j, j \neq k}} \frac{\mathrm{d}\alpha_k}{\mathrm{d}T} \right] \Delta T \right\} \tag{1.4}$$

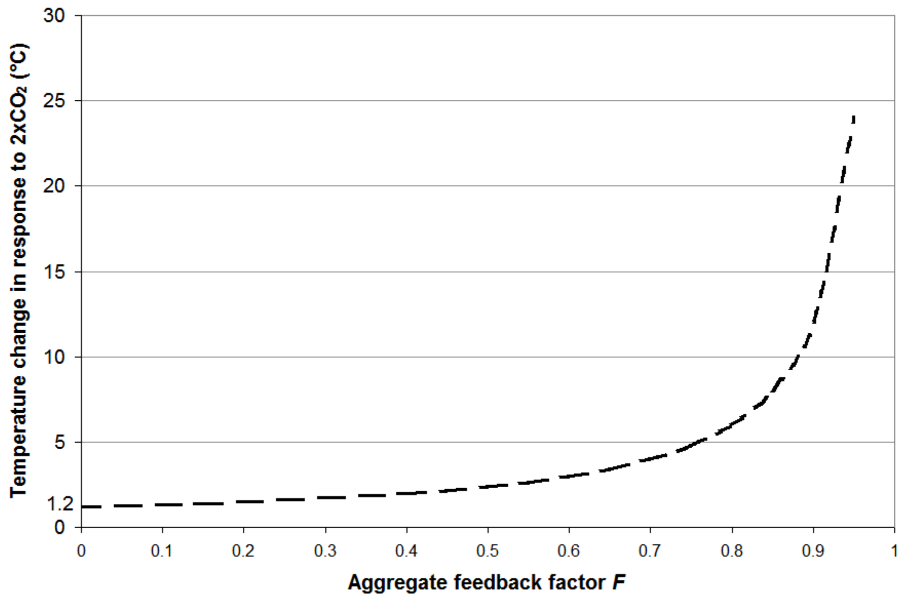


Figure 1.1: Temperature increases nonlinearly in the aggregate feedback factor F, as in equation (1.2). Shows the increase in temperature in response to doubled CO_2 concentrations ($\Delta R_f = 3.7$ W m^{-2} as in Forster et al., 2007).

Rearranging yields equation (1.2) again, where the feedback factors f_k are now defined as follows:

$$f_k = \lambda_0 \left\{ \left(\frac{\partial R}{\partial \alpha_k} \right)_{\alpha_{j,j \neq k}} \frac{\mathrm{d}\alpha_k}{\mathrm{d}T} \right\} \tag{1.5}$$

Each f_k (and F as well) is a non-dimensional measure that, when positive, can be interpreted as the fraction of the complete earth system warming due to that particular feedback process. If $F \geq 1$, then feedbacks would be responsible for 100% or more of the complete system warming, which is a nonsensical result indicating problems with the feedback model due to misunderstanding the reference system and feedback processes or due to the omission of countervailing negative feedbacks.

The present application of feedbacks potentially suffers from three complicating features: timescales, nonlinearities, and interactions. First, feedbacks differ in their effects over different timescales and in their speed. Water vapor feedback, for instance, may be negative over short timescales but positive over longer timescales (Hallegatte et al., 2006). Differences in speed matter for transient climate change, but this study looks only at equilibrium climate change. Second, feedbacks are often assumed to be linear, but they are clearly nonlinear over some sufficiently large range (e.g., Colman et al., 1997; Colman and McAvaney, 2009). In principle, nonlinearities can be included by deriving feedbacks via a second-order Taylor expansion (e.g., Roe, 2009). Third, the crucial assumption from, for instance, equations (1.1) and (1.3) that feedbacks are independent requires that the only way in which feedback processes interact is through surface temperature change. However, water vapor interacts with the lapse rate, and clouds interact with water vapor, surface albedo, soil moisture, and the lapse rate (Stephens, 2005; Bony et al., 2006). These interactions can be included by modifying equation (1.3) to allow nonzero cross-partials in a second-order Taylor series expansion of $\mathrm{d}R_\alpha$. Bony et al. (2006) suggested that nonlinearities and interactions may not be significant for moderate climate change, but because the current paper's results are meant to represent the possibility of extreme climate change, nonlinearities and interactions may provide opportunities for future extensions.

1.4 Methods: Hierarchical Bayes framework

Hierarchical methods use multilevel modeling to connect data to each other and to parameters of interest through distributions controlled by parameters drawn from their own distributions. In a Bayesian setting, some distributions are prior distributions reflecting beliefs about parameters formed before obtaining data and which are updated in response to data to form posterior distributions (e.g., Gelman et al., 2004). Hierarchical methods have been used in the climate science literature to connect data over varying spatial scales (e.g., Min and Hense, 2007), to consider optimal superensemble design and the development of climate forecasts (Berliner and Kim, 2008), and to include the possibility of shared structural biases in the course of developing a joint distribution for changes in temperature and precipitation (Tebaldi and Sansó, 2009). In the current model, the different levels of the hierarchy represent different types of uncertainty affecting estimation of feedback factors.

I define a "study" of a feedback factor to be a single climate model or empirical analysis. Each study can report more than one value ("observation") for the feedback factor, whether because of multiple ways of calculating the feedback factor (e.g., the use of different radiative kernels within a single climate model) or because of data-driven uncertainty (e.g., standard errors in empirical estimation). Each study's observations may cluster around a feedback value that is offset from the true value as a result of the study's biases, and each study may share more biases with some studies and less with others. We can therefore imagine a hierarchy of groups, with group membership determining how closely related the studies may be. One convenient framework divides studies into two groups: climate models and empirical studies that use climatic records. We might expect empirical studies of feedbacks to share biases if they come from time periods with different conditions than the present (e.g., more or less extensive land ice sheets), if they share measurement or dating errors, and if past climatic variability does not perfectly approximate the changes produced by anthropogenic GHG forcing. Climate models, in turn, also have several possible sources of shared biases (Tebaldi and Knutti, 2007): common resolution, common parameterizations or unresolved processes, shared observations used to tune the models, shared grids and numerical methods, shared components, and shared creators. Jun et al. (2008) provided evidence from late 20th century temperature simulations that models do have shared biases and even that models created by the same institution give more highly correlated output (see also Knutti et al., 2010). Each study therefore has uncertainty about the estimate it produces and about how its estimation procedure tends to produce results that differ from the values that will explain future climate change.

Figure 1.2 illustrates the model as applied in the current paper. Consider feedback factor f_k, out of K feedbacks. Let f_K represent all unknown and unmodeled feedbacks, so we have observations for only those feedbacks f_k with $k \in [1, K-1]$. Define a group j composed of studies i of feedback factor k such that, conditional on the true value of f_k, the group's study results M_{ijk} are exchangeable and so can be treated as if they come from a distribution over which we have prior beliefs (Bernardo and Smith, 1994). Exchangeability in this case means that all prior information about a study's outcomes is given by its group membership. I treat the study results as generated by a process with a normally distributed error term:

$$M_{ijk} \sim N(f_k + \theta_{jk}, \sigma^2_{jk}) \tag{1.6}$$

The studies in a group have a mean value that is offset from the true value f_k by an unknown amount θ_{jk} that represents the bias common to all members of group j.[3,4] Future applications may specify a higher-level distribution for θ_{jk} if groups may share some structural biases.

As described above, each study M_{ijk} in a group may report not just one point estimate for the feedback factor but either a distribution of values or a set of values. M_{ijk} can then be interpreted as the study's actual representation of feedback k, which we only observe with noise. If standard errors $\tilde{z}_{ijk}$ are available for each study observation $\hat{z}_{ijk}$, then assuming that $\hat{z}_{ijk}$ is a normally distributed unbiased estimator of M_{ijk} with df degrees of freedom yields:

$$\hat{z}_{ijk} \sim t(M_{ijk}, \tilde{z}_{ijk}, df) \tag{1.7}$$

where $t(x, y, z)$ is a t distribution with location parameter x, scale parameter y, and shape parameter z. Alternately, as is the case throughout this application, study M_{ijk} may report several values y_{hijk}. These observations may be combined in lower-level groups used, inter alia, to inform the parameters for M_{ijk}. I assume the within-study observations are normally distributed with mean M_{ijk}:

$$y_{hijk} \sim N(M_{ijk}, \phi^2_{jk}) \tag{1.8}$$

[3] As currently implemented, the statistical framework treats all data sources as equally reliable. Tebaldi and Knutti (2007) and Tebaldi and Sansó (2009) reviewed approaches that have used observational data to weight climate models in an ensemble or to weight combinations of parameters within a single model, often assuming that the optimal weighting does not change from the calibration data set to the future. Knutti et al. (2010) described some of the conceptual difficulties in determining the optimal weighting.

[4] In representing models as clustering around a common offset from f_k, the proposed methods take advantage of the fact that the most complex models do not sample the range of uncertainty but are calibrated to give their best estimates (Knutti, 2008; Knutti et al., 2010).

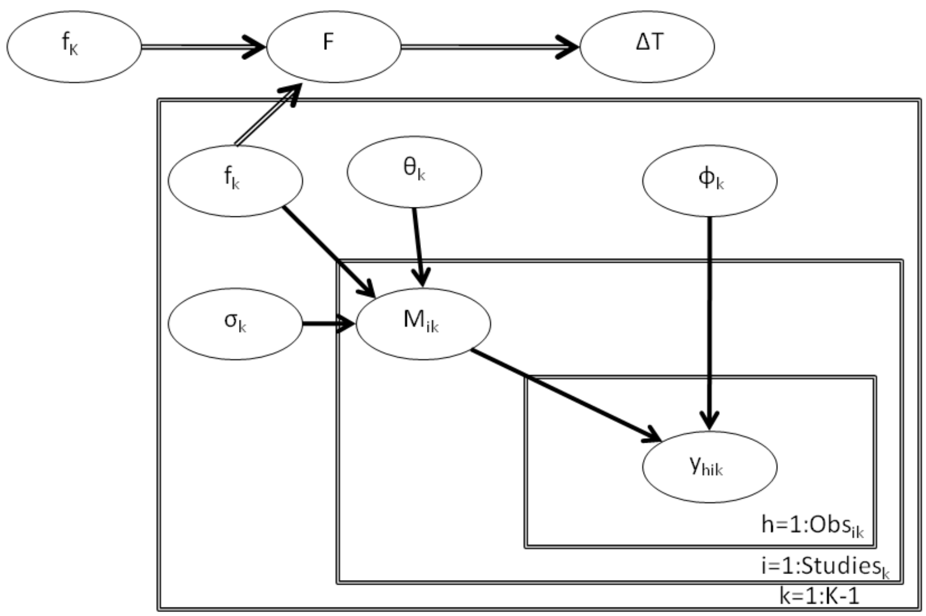


Figure 1.2: A representation of the statistical model following the conventions for a WinBUGS directed acyclic graph (Lunn et al., 2000). The j subscripts are omitted because the current application only has one group per feedback factor. The solid arrows indicate stochastic dependence, and the hollow arrows indicate deterministic dependence.

The intra-study standard deviations ϕ_{jk} are here shared among studies in a group because the current application's intra-study variation comes from combining the same three radiative kernels with a given climate model's output. In other applications, the within-study standard deviation could be specific to each study i or could be drawn from a group-level distribution.

This hierarchical approach clearly separates several sources of error or bias: 1) the structural bias common across groups is the expected value of any higher-level distribution that may exist for θ_{jk}, 2) the structural bias specific to group j is the difference between θ_{jk} and the structural bias common across groups or, if θ_{jk} lacks a higher-level distribution, is θ_{jk}, 3) the between-study variation within a group is determined by σ_{jk}, 4) the uncertainty within a particular study's observations due to limited data, sampling variability, or different ways of obtaining data is determined by $\tilde{z}_{ijk}$ or ϕ_{jk}, and 5) the possibility of omitted feedbacks is represented by a feedback term f_K lacking observations.

Applying Bayes' Theorem updates prior distributions in response to observations of feedback factors to produce posterior distributions for all parameters. Table 1.1 describes the six sets of priors used for the standard deviations of within-study observations (ϕ_{jk}), the between-study standard deviations (σ_{jk}), the estimated feedback factors (f_k), the unknown feedback factor (f_K), and the biases shared between a group's studies (θ_{jk}). Each prior distribution is independent of all others. Figure 1.3 plots each type of prior distribution used. Care must be taken in prior selection (e.g., Frame et al., 2005), as flat (uniform) priors on one type of parameter can contribute more information to another parameter than intended (Dongen, 2006). Except where varied to assess sensitivity, we use weakly informative priors that concentrate prior probability mass in the range of the a priori most plausible values (i.e., values closer to 0) while still placing significant probability on more extreme values. These weakly informative priors aim to let even sparse data dominate the posterior distributions without ruling out extreme values. The use of weakly informative priors can be important because with small numbers of groups and, possibly, of observations within a group, less informative priors may not be dominated by the data and may produce results that are highly sensitive to the form of the prior (Kass and Wasserman, 1996; Lambert et al., 2005).[5,6] The specific numerical values for the priors are meant to give reasonable-looking distributions, and the six sets of priors will help assess the sensitivity of the posterior distributions to the form of the prior.

Annan and Hargreaves (2011) argued that previous analyses' use of a uniform prior for climate sensitivity made their results sensitive to the prior's upper bound and generally led to overly pessimistic conclusions about the possibility of extreme temperature change. The prior distributions on f_k and f_K imply prior distributions for climate sensitivity that are

[5] Interestingly, sensitivity to prior beliefs may partially explain the actual diversity of posterior beliefs about climate change.

[6] Gelman (2006) explored the choice of priors for between-group variance parameters when group size is small. For cases with at least 5 groups, he recommended using a uniform reference prior on the standard deviation, and for cases with fewer groups, he recommended the folded non-central t distribution, of which the half-Cauchy used here is a special case.

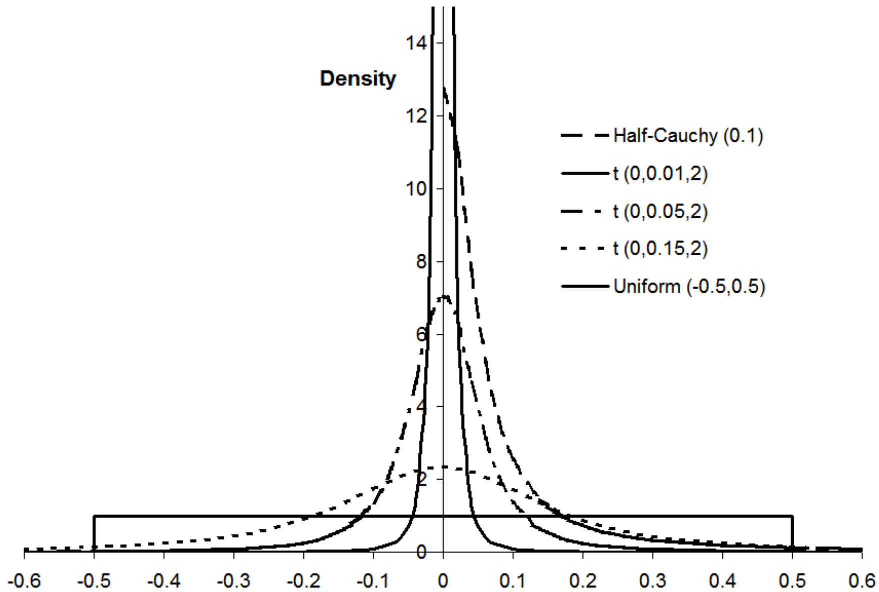


Figure 1.3: The five types of prior distributions used in this application of the model, as described in Table 1.1. $\mathrm{t}(x, y, z)$ refers to a t distribution with location parameter x, scale parameter y, and shape parameter z. The t distribution with scale parameter 0.01 is only used in the learning experiment described in section 7.

far from uniform and that concentrate prior probability in the IPCC's likely range, yet we will see that these prior distributions nonetheless can generate posterior distributions with significant probability of extreme temperature change.[7]

The two most significant omissions in the proposed statistical model are constraints on the total system response and the possibility of abrupt changes and threshold effects. First, observations of, for instance, temperature change and ocean heat capacity may be able to constrain F directly even though these do not provide information about any particular feedback process. These observations may therefore be able to constrain the positive tail of the distribution for ΔT, but they will likely have their own structural biases and are unlikely to account for the operation of slower feedbacks (Urban and Keller, 2009). Second, concern about climate change may be driven not just by concern about its ultimate magnitude but also by concern about its possibly being both abrupt and irreversible (Keller et al., 2008; Lenton et al., 2008; Solomon et al., 2009). Both abruptness and irreversibility should be included in a complete risk assessment and should inform near-term abatement decisions.

The posterior distributions were sampled using Markov chain Monte Carlo methods as implemented in WinBUGS version 1.4.3 (Lunn et al., 2000; Gamerman and Lopes, 2006).

[7]In addition, the implied prior distributions for climate sensitivity do not differ greatly from one prior combination to another.

Table 1.1: The six combinations of prior distributions for model parameters (also see Figure 1.3). HC(x) is a half-Cauchy distribution with scale parameter x, U(x, y) is a uniform distribution on (x, y), and t(x, y, z) is a t distribution with location parameter x, scale parameter y, and shape parameter z.

Prior	f_k [a,b]	f_K [a,b]	θ_k [c]	ϕ_k	σ_k
1	t(0,0.15,2)	t(0,0.15,2)	t(0,0.05,2)	HC(0.1)	HC(0.1)
2	U(-0.5,0.5)	t(0,0.15,2)	t(0,0.05,2)	HC(0.1)	HC(0.1)
3	t(0,0.15,2)	t(0,0.05,2)	t(0,0.05,2)	HC(0.1)	HC(0.1)
4	t(0,0.15,2)	t(0,0.05,2)	Pr(θ_{jk}=0)=1	HC(0.1)	HC(0.1)
5	t(0,0.15,2)	Pr(f_K=0)=1	Pr(θ_{jk}=0)=1	HC(0.1)	HC(0.1)
6	U(-0.5,0.5)	Pr(f_K=0)=1	Pr(θ_{jk}=0)=1	HC(0.1)	HC(0.1)

[a] The t distributions are censored so that values are observed to be less than 1.

[b] F is constrained to be less than 1.

[c] The t distributions are censored so that values are observed to be between -0.5 and 0.5.

Each posterior distribution generated one million samples after a burn-in period of one million samples. The sample size was large enough for multiple chains to converge on the posterior distributions.

1.5 Data: Models' estimates of feedbacks

This paper considers four climate change feedbacks for which model estimates are available: albedo, clouds, water vapor-lapse rate, and carbon cycle. It also considers the sum of all other feedbacks, including unmodeled and unknown feedbacks. The water vapor and lapse rate feedbacks can be treated as a combined water vapor-lapse rate feedback because of their strong negative correlation noted by, among others, Bony et al. (2006), Soden and Held (2006), and Soden et al. (2008). All calculations in this paper use $\lambda_0 = 0.315$ K (W m^{-2})$^{-1}$ because Soden et al. (2008) found that λ_0 ranges roughly from 0.31 to 0.32 K (W m^{-2})$^{-1}$.

Methods for calculating feedback factors from climate models often rely on representations like equation (1.5). Soden et al. (2008) elaborated a method that can enable consistent comparison between models while avoiding biases caused by correlation between climate fields. Their method decomposes feedbacks into the mean change in the associated climate field as the climate is perturbed and a radiative kernel that gives the change in radiative forcing due to a change in the climate field. The kernel is independent of the general circulation model for which the feedback is calculated and depends only on the control climate. Soden et al. (2008) used three different radiative kernels to estimate albedo, cloud, and water vapor-lapse rate feedback factors in each of the climate models from the Intergovernmental Panel on Climate Change's Fourth Assessment Report that provided the necessary data. They updated and improved the earlier analyses of Colman (2003), Winton (2006), and Soden and Held (2006). Figure 1.4 and Table 1.2 show the data from Soden et al. (2008) using their model acronyms and converted to the non-dimensional feedback form. As previously reported (e.g., Bony et al., 2006), variance in cloud feedback estimates is primarily responsible for the variance in estimates of the aggregate feedback factor

Ideally, observations based on climatic records could supplement these observations from climate models so as to obtain a group of observations that does not share the structural biases common across models. However, several hurdles deter inclusion of empirical estimates in this paper. First, some of the empirical observations have been implicitly accounted for in model development. Second, it is difficult to compute the partial derivatives in equation (1.5) in a way that ensures that only one variable is changing (Bony et al., 2006), climatic variability can create bias in estimations (Spencer and Braswell, 2008), and results may need to be adjusted to be relevant to the present and future (Yoshimori et al., 2009). In one illustration of the potential complications, de F. Forster and Collins (2004) and Dessler et al. (2008) attempted to estimate the water vapor feedback from responses to volcanic and El Niño forcings, but these forcings may not adequately mimic long-term climate change (Bony et al., 2006; Dessler et al., 2008) and the data and methods may not match the the timescales of relevant processes (see Hallegatte et al., 2006). Carefully sorting and improving the empirical literature could contribute to extending the present paper's results.

Carbon cycle feedbacks include processes by which temperature changes alter atmospheric CO_2 concentrations, which in turn affect radiative forcing and so temperature. A common definition of carbon cycle feedbacks includes warming-induced changes in CO_2 sources or sinks but does not include changes in CO_2 sources or sinks due directly to changing CO_2 concentrations (e.g., Friedlingstein et al., 2006).[8] This definition implies that the aggregate feedback factor F applies to the radiative forcing resulting from the CO_2 concentrations obtained after combining a CO_2 emission profile with an offline model of how the carbon cycle responds to increased CO_2 levels, holding climate constant. In other words, the aggregate feedback factor applies to a CO_2 concentration already adjusted for CO_2 fertilization and (non-biologically) changing ocean sinks. Vegetative and oceanic processes can be net sinks even if the carbon cycle feedback is positive because positive feedback here just means that the strength of these sinks decreases with warming.

Data for carbon cycle feedbacks come from the 11 models of the Coupled Climate-Carbon Cycle Model Intercomparison Project (C^4MIP), as reported by Friedlingstein et al. (2006)

[8] More recent work by Gregory et al. (2009) formally separated carbon cycle feedbacks into concentration-carbon feedbacks and climate-carbon feedbacks.

and as adjusted by Cadule et al. (2009) for the nonlinearity of radiative forcing as a function of CO_2 levels.[9] However, as shown in the appendix, both analyses implicitly included uncoupled models' feedbacks in their reference systems. With regard to the present purposes, they therefore overestimated carbon cycle feedbacks by including the operation of albedo, water vapor, lapse rate, and cloud feedbacks in response to carbon cycle feedbacks' effects on CO_2 concentrations. To then obtain an estimate of the aggregate feedback factor F by summing their estimates of the carbon cycle feedback with estimates of these other feedbacks would double-count the other feedbacks. The appendix explains how to adapt these carbon cycle feedback estimates, and Figure 1.4 and Table 1.2 show their adjusted values.

Models' estimates of carbon cycle feedback factors are notably incomplete because most only represented changes in photosynthesis, growth, and decomposition. Possibly important processes largely omitted by the models considered in Friedlingstein et al. (2006) include permafrost melting, fires, tropical deforestation, and nutrient limitation (Field et al., 2007; Schuur et al., 2008; Sokolov et al., 2008; Schuur et al., 2009; Tarnocai et al., 2009),[10] though more are including dynamic vegetation that allows for phenomena such as biome shifts and fires (Field et al., 2007). Further, the models do not fully explore the parameter space for the processes they do include (Matthews et al., 2007), and their parameterizations probably share biases. Most models report positive net carbon cycle feedbacks because of "the stimulated net C release from land ecosystems in response to climate warming" (Luo, 2007: p. 687). However, while this carbon release is primarily driven by models' similar representations of the sensitivity of photosynthesis and respiration to changing temperatures, there is still much uncertainty about the direction and degree of these responses (Luo, 2007).

The difficulty of modeling the carbon cycle makes estimation from climatic records especially desirable, but it is also difficult to empirically estimate carbon cycle feedbacks from past observations because CO_2 levels affect temperature even as temperature affects CO_2 levels. This results in a system of simultaneous equations, with the noise term in one equation correlated with the other equation's dependent variable and so with its own regressor. Because of this correlation between regressor and noise, ordinary least squares regression produces inconsistent estimates. Studies by Scheffer et al. (2006) and Torn and Harte (2006) assumed that there were no significant non-temperature drivers of CO_2 concentrations in the last millennium's Little Ice Age or in a 250,000 year portion of the Vostok ice core record, which would avoid asymptotic bias by eliminating the noise term in the regression of CO_2 on temperature. However, because even small biases in feedback estimates can have special significance due to the nonlinearity of temperature change in the aggregate feedback factor F, this assumption requires further validation. These studies also did not report standard errors from their regressions, making it difficult to place their results in a probabilistic analysis. Lemoine (2010b) uses orbital forcing to instrument for temperature change, attempting

[9]These models provide a transient feedback analysis, not an equilibrium analysis. This is one shared source of bias when used for estimates of the equilibrium feedback factor.

[10]Note that the carbon cycle feedback could be defined so that some of these were different feedbacks.

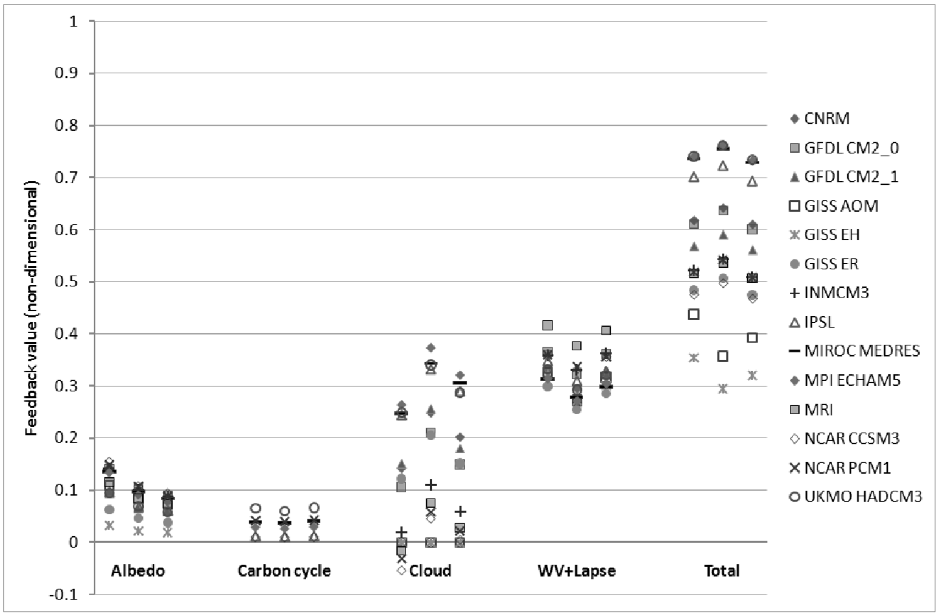


Figure 1.4: The data used as input to the statistical model (see also Table 1.2), plus a column showing the sum of the feedback factors for each model. Each feedback factor's observations are grouped by the radiative kernel used to calculate them from the indicated general circulation model. The BMRC kernel is on the left, the GFDL kernel is in the middle, and the NCAR kernel is on the right. Data are from Soden et al. (2008) and Cadule et al. (2009), with the latter adjusted as described in the appendix. Model and kernel names follow Soden et al. (2008).

to obtain unbiased estimates of feedback strength. Properly calibrating this uncertainty is important because, as shown below, having a group of studies not sharing models' biases can greatly affect the posterior distributions for feedback strength and temperature.

One strength of the hierarchical Bayes framework is its ability to assess the importance of beliefs about unknown and unmodeled feedbacks. Climate models do not include all feedbacks. Known missing feedbacks include changes in non-CO_2 GHGs, in subsea methane hydrates, in ocean circulation and biota, in ocean sinks via changing wind regimes, in albedo and other biophysical feedbacks due to ecosystem responses, and, for transient climate change, in sea surface temperatures due to hydrological cycle intensification (Gruber et al., 2004; Archer, 2007; Field et al., 2007; Le Quéré et al., 2007; Williams et al., 2007; Lawrence et al., 2008; Cadule et al., 2009; Dorrepaal et al., 2009; Zeebe et al., 2009). Furthermore, slow feedbacks such as land ice sheet dynamics that become significant after decades or centuries of forcing should be considered in discussions of GHG stabilization targets (Hansen et al., 2008). The prior on f_K should represent uncertainty about the aggregate strength of omitted feedbacks.

Table 1.2: The non-dimensional feedback factors calculated from Soden et al. (2008) and Cadule et al. (2009), as described in the text and in the appendix. Model and kernel names follow Soden et al. (2008). Figure 1.4 plots these data.

Model	Radiative Kernel	Albedo	Carbon Cycle	Cloud	WV+Lapse
	BMRC	0.14		0.14	0.33
CNRM	GFDL	0.10		0.25	0.29
	NCAR	0.087		0.20	0.32
	BMRC	0.14		0.11	0.36
GFDL_CM2_0	GFDL	0.10		0.21	0.32
	NCAR	0.089		0.15	0.36
	BMRC	0.097		0.15	0.32
GFDL_CM2_1	GFDL	0.065		0.26	0.27
	NCAR	0.059		0.18	0.32
	BMRC	0.11			0.33
GISS_AOM	GFDL	0.085			0.27
	NCAR	0.075			0.32
	BMRC	0.032			0.32
GISS EH	GFDL	0.022			0.27
	NCAR	0.019			0.30
	BMRC	0.063		0.12	0.30
GISS ER	GFDL	0.047		0.21	0.25
	NCAR	0.038		0.15	0.29
	BMRC	0.14		0.019	0.36
IMMCM3	GFDL	0.10		0.11	0.33
	NCAR	0.087		0.059	0.36
	BMRC	0.096	0.012	0.24	0.35
IPSL	GFDL	0.069	0.012	0.33	0.31
	NCAR	0.061	0.013	0.29	0.33
	BMRC	0.14	0.039	0.25	0.31
MIROC MEDRES	GFDL	0.097	0.037	0.34	0.28
	NCAR	0.084	0.041	0.31	0.30
	BMRC	0.13	0.029	0.26	0.31
MPI ECHAM5	GFDL	0.091	0.026	0.37	0.27
	NCAR	0.078	0.030	0.32	0.30

Continued on next page

Table 1.2 – continued from previous page

Model	Radiative Kernel	Albedo	Carbon Cycle	Cloud	WV+Lapse
	BMRC	0.12		-0.016	0.42
MRI	GFDL	0.084		0.075	0.38
	NCAR	0.073		0.028	0.41
	BMRC	0.16	0.011	-0.053	0.36
NCAR CCSM3	GFDL	0.11	0.010	0.056	0.33
	NCAR	0.095	0.011	0.0056	0.36
	BMRC	0.15	0.042	-0.031	0.36
NCAR PCM1	GFDL	0.11	0.040	0.058	0.34
	NCAR	0.088	0.043	0.021	0.36
	BMRC	0.095	0.065	0.25	0.33
UKMO HADCM3	GFDL	0.069	0.060	0.34	0.29
	NCAR	0.058	0.067	0.29	0.32

1.6 Results: Posterior distributions

Posterior distributions for the albedo, carbon cycle, cloud, and water vapor-lapse rate feedbacks result from using the data given in Table 1.2 and Figure 1.4 to update each of the six combinations of prior distributions from Table 1.1 within the hierarchical Bayes framework shown in Figure 1.2. These posterior distributions combine with the posterior distribution for the unknown and unmodeled feedbacks f_K (which is the same as its prior distribution in Table 1.1 due to the absence of data) to produce posterior distributions for the aggregate feedback factor F and for the temperature change ΔT_{2xCO2} resulting from doubling CO_2 concentrations (where the CO_2 doubling does not account for carbon cycle feedbacks).[11]

Figures 1.5 and 1.6 show the posterior distributions resulting from each set of priors. Figure 1.7 gives the same results in box plot form. Prior combinations 5 and 6 vary the prior on f_k when there is no possibility of shared structural bias, showing that the posterior for each feedback factor f_k is not sensitive to the form of the prior on f_k in the certain absence of shared structural biases. However, comparing prior combinations 1 and 2 shows that the prior for f_k can affect the posterior when there is a nonzero probability of shared structural bias. Prior combinations 1 and 3 only vary the prior on f_K, and the same applies to prior combinations 4 and 5. Comparing the posterior distributions produced within each

[11]Doubling CO_2 concentrations produces additional radiative forcing ΔR_f of 3.7 W m^{-2} (Forster et al., 2007: 140).

pair of priors shows how the unknown and unmodeled feedbacks f_K spread the probability distribution for temperature change, especially on the high side. Because the prior on f_K is never updated, it is directly combined with the posteriors for the four constrained feedbacks to obtain the posterior for F, and because the possibility of slightly greater values of net positive F spreads the temperature change distribution more than does the possibility of slightly lower values (Roe and Baker, 2007; Hannart et al., 2009), the nonzero priors on f_K thicken the positive tail of the temperature change distribution more than they thicken the negative tail.

The posterior distributions are sensitive to the possibility that models share biases (compare prior combinations 3 and 4). Possible shared biases have three effects. First, they increase the spread of the posterior distribution for each f_k. The possibility of $\theta_{jk} < 0$ makes relatively high values of f_k more probable, and the possibility of $\theta_{jk} > 0$ makes relatively low values of f_k more probable. Second, the possibility of shared biases limits the benefit of additional models because, with only one group of observations per feedback factor, the available observations can only constrain the sum of the feedback factor and the shared bias term ($f_k + \theta_{jk}$). Even with an unbounded number of models, we could never be sure what portion of their signal is related to the true feedback and what portion is related to shared biases.

Third, the possibility of shared biases can shift the posterior median for f_k towards 0 when the prior distribution for f_k has a peak at 0. The posterior distributions for f_k and θ_{jk} depend on their prior distributions and on the posterior for $f_k + \theta_{jk}$ (Figure 1.8). When both priors are t distributions with peaks at 0 (as in prior combinations 1 and 3), both posteriors move towards the posterior for $f_k + \theta_{jk}$; however, neither parameter's posterior is identical to the posterior for $f_k + \theta_{jk}$ because values with greater prior probability for one parameter (e.g., f_k closer to 0) often imply values with lower prior probability for the other (e.g., implying θ_{jk} farther from 0). For net positive $f_k + \theta_{jk}$, the posterior for θ_{jk} will peak at a positive value and the posterior for f_k will peak at a value less than the peak of $f_k + \theta_{jk}$. In contrast, when the prior on f_k is a uniform distribution (as in prior combination 2), the posterior on θ_{jk} is approximately the same as its prior because (ignoring effects from the boundary of the uniform distribution) the posterior on f_k can take any form without sacrificing prior knowledge. Each value for f_k has equal prior probability over the range of the uniform distribution, which allows the posterior for θ_{jk} to be completely determined by its prior knowledge. In order to keep the posterior for θ_{jk} in its region of greatest prior probability, less peaked priors on f_k therefore tend to center the posterior for f_k on the posterior for $f_k + \theta_{jk}$.

From this point forward, prior combination 3 is the base case because it includes the possibility of shared structural biases, includes the possibility of unmodeled and unknown feedbacks, and uses a prior on the constrained feedbacks that makes small values the most likely while allowing the possibility of extreme outcomes. Without clear grounds for favoring either prior 1 or prior 3's representation of omitted feedbacks, I use prior 3 only to show that the following results are not driven by a highly diffuse prior on omitted feedbacks. Prior

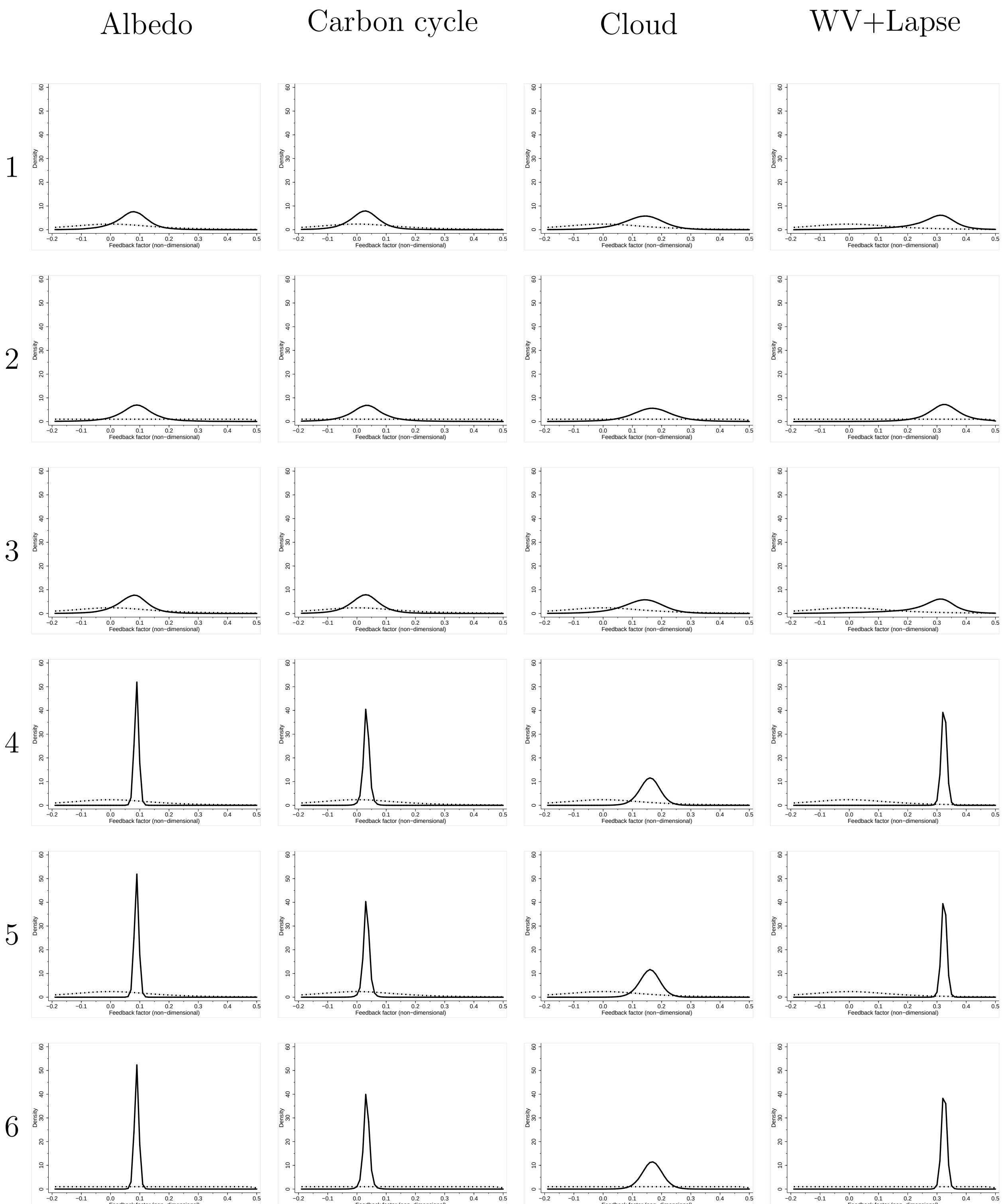


Figure 1.5: Prior distributions (dashed) and posterior distributions (solid) for the four feedbacks f_k (columns) under the 6 combinations of priors (rows). Prior combinations 1 through 3 place positive probability on nonzero shared structural bias, and prior combinations 1 through 4 place positive probability on models being incomplete.

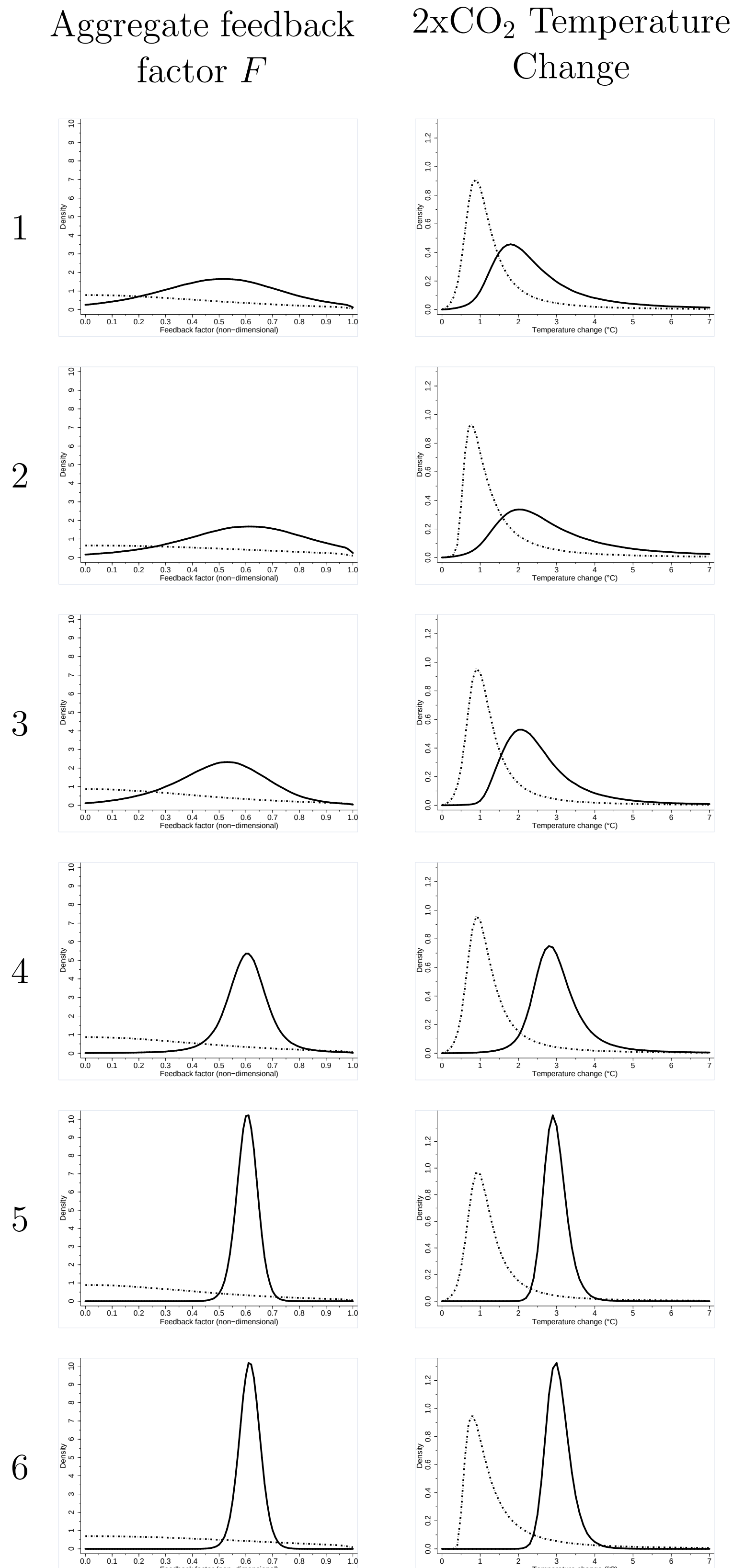


Figure 1.6: Prior distributions (dashed) and posterior distributions (solid) for the aggregate feedback factor F and for the temperature change in response to doubled CO_2 concentrations (relative to pre-industrial levels and with the doubling measured before accounting for carbon cycle feedbacks) under the 6 combinations of priors (rows).

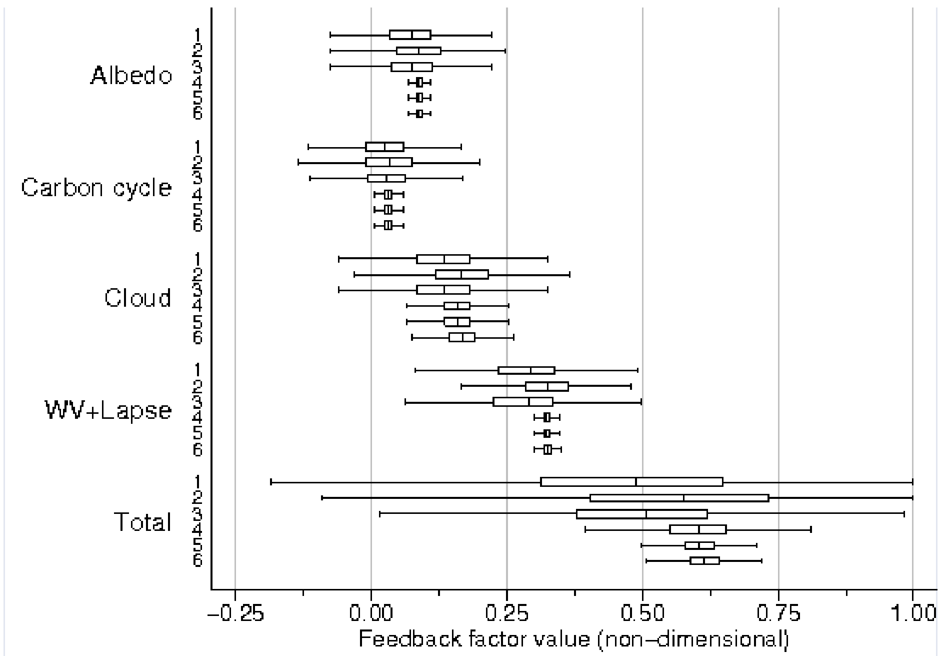
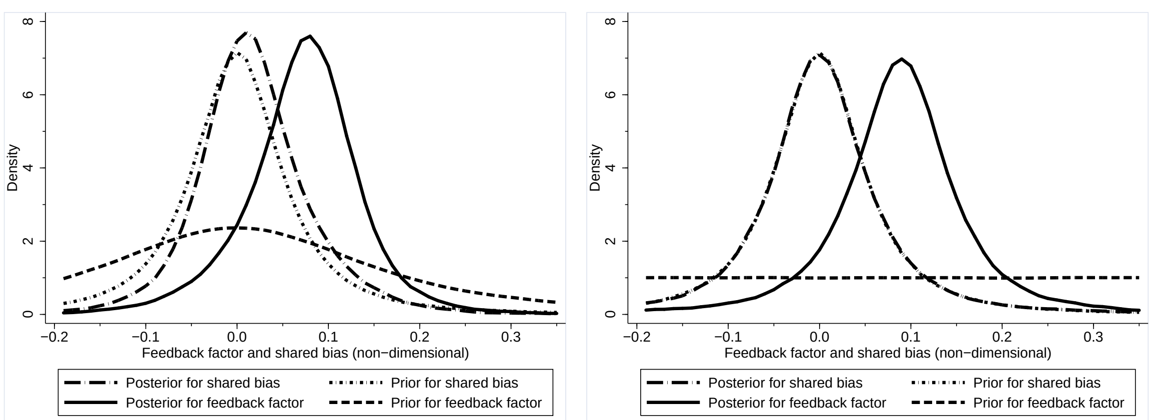


Figure 1.7: Box plots for the posterior distributions of the individual feedback factors f_k and the aggregate feedback factor F under each of the 6 combinations of priors. The boxes show the interquartile range and median value, and the whiskers define the adjacent values. The plot does not show outside values and, for computational convenience, uses only the first 500,000 samples from the distributions in Figures 1.5 and 1.6.



(a) From prior combination 1

(b) From prior combination 2

Figure 1.8: The posterior and prior distributions for the shared bias term (θ_{jk}) and feedback factor (f_k) for the albedo feedback. Shows prior combinations 1 and 2, which vary the form of the prior on f_k. With prior combination 2, the prior and posterior for θ_{jk} completely overlap.

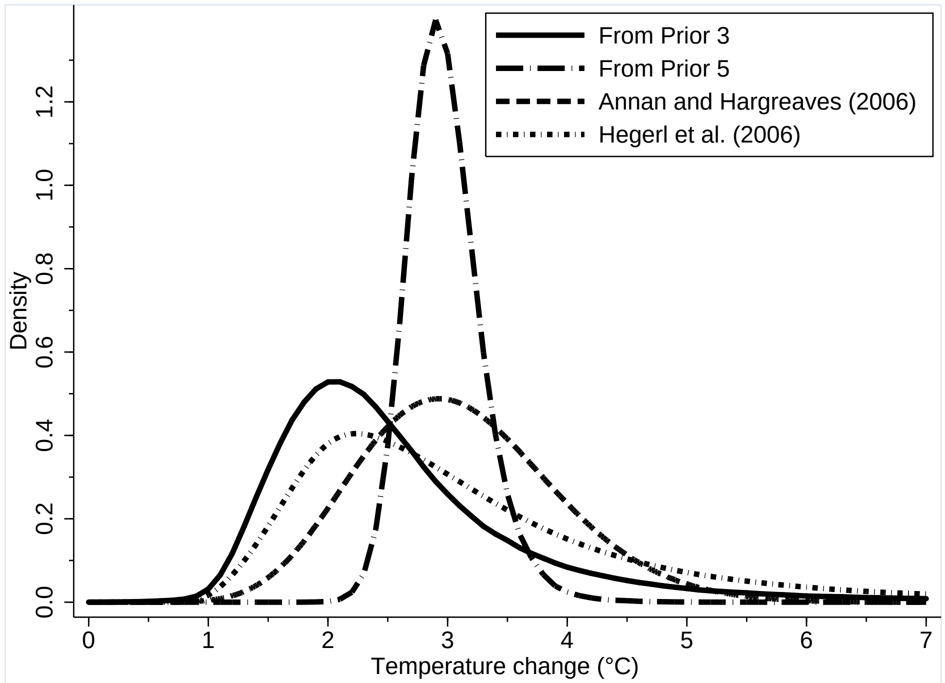


Figure 1.9: Posterior probability distributions for the temperature change in response to doubled CO_2 concentrations (relative to pre-industrial levels) for prior combination 3 (which allows nonzero shared bias and omitted feedbacks) and for prior combination 5 (which assumes models are independent and complete). Also, the distribution from Annan and Hargreaves (2006) using the combination of three constraints and the distribution from Hegerl et al. (2006) using all pre-industrial proxies with a prior derived from the instrumental record.

combination 5 provides an important, if unrealistic, reference case because it represents maximal confidence in climate models' completeness and independence. Figure 1.9 compares the posterior distributions resulting from these two prior combinations with the distributions for climate change produced by Annan and Hargreaves (2006) and Hegerl et al. (2006) using multiple constraints from the instrumental and paleoclimatic records. Prior combination 5 generates the most peaked posterior and the posterior with the peak at the greatest temperature change. The other distributions indicate the greatest risk of extreme temperature change: relative to the results from prior combination 5, they shift probability mass from moderate temperature change to extreme temperature change. Prior combination 3 carries this trend the farthest (though still not as far as prior combinations 1 and 2), placing the most probability mass both on low temperature change and on high temperature change because the possibility of shared biases makes extreme values of each f_k more plausible.

An alternate means of generating probability distributions would identify each model's output with a single point estimate that is the mean of the three radiative kernel results and then treat the models as independent, identically distributed samples from a normally distributed population (similar to Roe and Baker, 2007). This method produces results nearly identical to those with prior combination 5, implying that the weakly informative prior distributions do not have a noticeable effect in the certain absence of shared structural biases. Importantly, therefore, the assumption of normally distributed feedbacks is not sufficient to generate distributions for temperature change with the long positive tail we have come to expect from past studies, partly because constraining uncertainty about feedbacks has a greater effect on temperature change distributions than Roe and Baker claimed (Hannart et al., 2009). If developing a distribution from knowledge of feedbacks, whether or not the distribution has the customary shape in the positive tail in fact depends strongly on whether the model includes the possibility of nonzero shared structural bias.

1.7 Discussion

1.7.1 Opportunities for learning

This statistical framework can readily incorporate new information to produce updated posterior distributions. Ignoring the possibility of modeling new feedback processes, we can learn in at least four ways:

1. By obtaining additional model observations (obtaining y_{hijk} for new i).
2. By obtaining observations that do not share the current observations' structural biases (adding a group j with its own M_{ijk}).
3. By decreasing the variance of the prior distribution for shared structural biases (constraining the prior for θ_{jk}).
4. By decreasing the variance of the prior distribution for unknown and unmodeled feedbacks (constraining the prior for f_K).

Figure 1.10 shows the posterior distributions resulting from prior combination 3 as well as from each of these four possible ways of learning. I represent obtaining more observations by including 6 more studies M_{ijk} for each f_k, where each new study contains three observations equal to the mean of the set of real observations. By reinforcing the actual observations' central tendency, these additional observations should constrain the posterior distribution at least as well as any other set with the same number of additional observations. I represent adding a group j by randomly assigning the existing models to one of two groups that are assumed not to share any structural biases. 10 models end up in a first group and 4 end up in a second, with each group including 3 models with carbon cycle feedback observations. The assumption of independence conditional on f_k and the similarity of the two groups' estimates together imply that the additional group should constrain the posterior distributions as much as possible given the group assignments. I represent decreasing uncertainty about shared

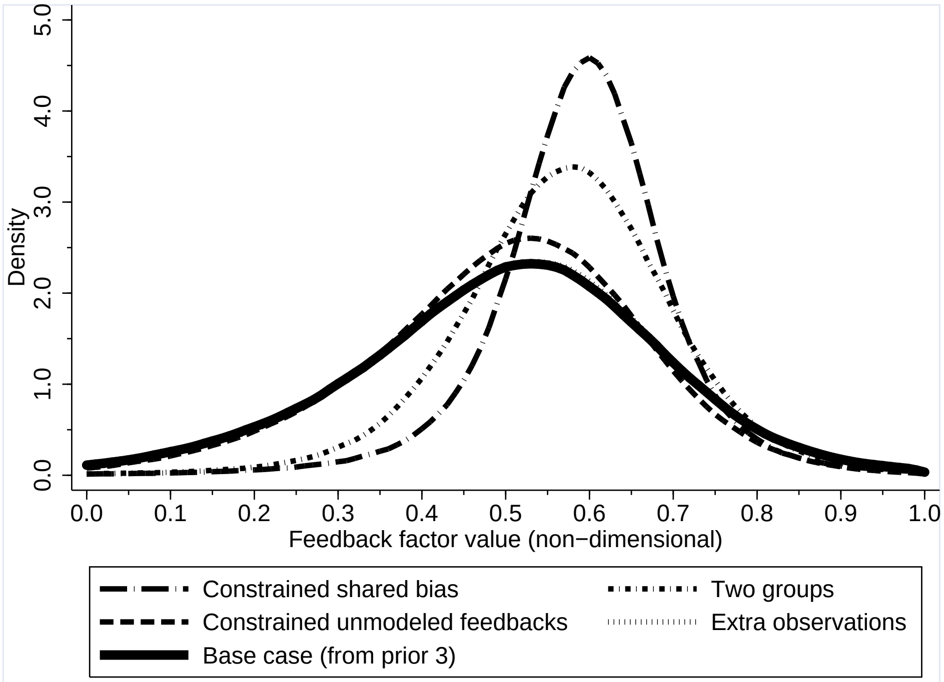


Figure 1.10: The posterior probability distribution for the aggregate feedback factor F after different types of hypothetical learning. The distribution with extra observations is not visible because it is almost exactly the same as the base case.

biases θ_{jk} and about unmodeled and unknown feedbacks f_K by replacing their priors with a t distribution having scale parameter 0.01, which is smaller than the scale parameter of 0.05 used for each in prior combination 3 (see Figure 1.3).

Remarkably, the additional observations do not perceptibly affect the posterior probability distribution. Obtaining additional observations of feedback factors from climate models related to existing ones only affects the posterior distributions if the new estimates differ from the current ones' central tendency. The present uncertainty is not driven by a lack of models. These results cohere with the findings in Figures 1.5 and 1.6 that the prior for f_k does not greatly influence the posterior distributions in the absence of possible shared biases.

Constraining unmodeled and unknown feedbacks does have a noticeable impact, but the impacts of obtaining an additional group and of constraining shared biases are more important. Having an additional, conditionally independent group enables each f_k to be constrained via two different $f_k + \theta_{jk}$ terms and so can have a similar, though less stark, effect to that of the simulated narrowing of prior beliefs about shared biases θ_{jk}. The additional group constrains the posterior distribution for the aggregate feedback factor F even though each group contains fewer observations than did the original group.[12] The

[12]The effects of a second group depend on how similar its estimates are to those of the first group.

simulated narrower prior beliefs about shared biases result in the narrowest probability distribution for the aggregate feedback factor F. Aside from explicitly modeling some of the currently unmodeled feedbacks (in which case the results would depend on how the newly modeled feedbacks affect beliefs about remaining unmodeled and unknown feedbacks), the activities that could most constrain the posterior probability distributions for the aggregate feedback factor and for temperature change are those that more tightly constrain prior beliefs about shared structural biases or those that produce observations of feedback factors through methods that are unlikely to share much structural bias with existing climate models.

1.7.2 Measures of temperature risk

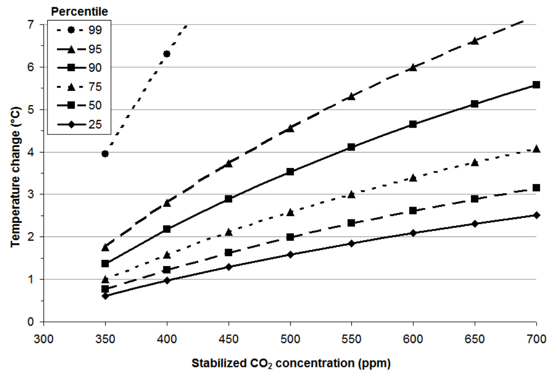
By focusing on factors that drive uncertainty about low-probability temperature change, the posterior distributions enable assessments of temperature change risks.[13] Table 1.3 gives percentiles and conditional expectations for the aggregate feedback factor F, and Figure 1.11 plots percentile temperature change against CO_2 concentrations.[14]. The percentile temperature change resulting from prior combination 3 is similar to those of Annan and Hargreaves (2006) and Hegerl et al. (2006) up to about the 95th percentile, at which point the value resulting from prior combination 3 increases faster because its posterior distribution has more weight in its extreme positive tail. Using prior combination 1 would amplify this result by placing even more weight in the posterior tails. If we believe the models cluster around the true outcomes and include all relevant processes (as in prior combination 5), then we have a distribution for climate sensitivity that places overwhelming probability on its being close to 3°C and almost no probability on its being greater than 4°C. However, including these other sources of uncertainty (as in prior combination 3) dramatically expands the positive tail so that there is more than a 5% chance that climate sensitivity is greater than 5°C. The expected climate sensitivity conditional on being above the 95th percentile is 18°C, and the expected climate sensitivity conditional on being above the 75th percentile is still 7°C. The temperature risk curves implied by Annan and Hargreaves (2006) and Hegerl et al. (2006) generally fall between the curves implied by prior combinations 3 and 5.

The CO_2 concentration needed to meet a 2°C target relative to pre-industrial levels depends strongly on risk tolerance and on prior beliefs about shared model biases and about model completeness.[15] If models are unrealistically assumed to be complete and to lack shared biases (prior combination 5), then CO_2 concentrations could stabilize at 410 ppm

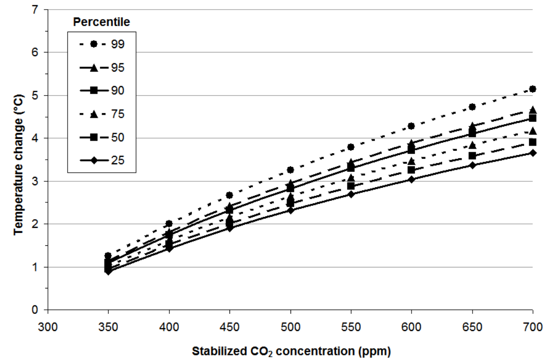
[13]We could define Temperature-at-Risk (TaR) and Conditional Temperature-at-Risk (CTaR) by analogy with the Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) metrics used to evaluate investment portfolios or to determine banks' capital requirements (Artzner et al., 1999; Rockafellar and Uryasev, 2000). VaR measures often use the 90th, 95th, and 99th percentiles (Rockafellar and Uryasev, 2000).

[14]Figure 1.11 calculates the temperature change for a CO_2 concentration C by using equation (1.2) and the relation $\Delta R_f = 5.35 \ln(\frac{C}{C_0})$, where C_0 is the pre-industrial concentration (e.g., Cadule et al., 2009)

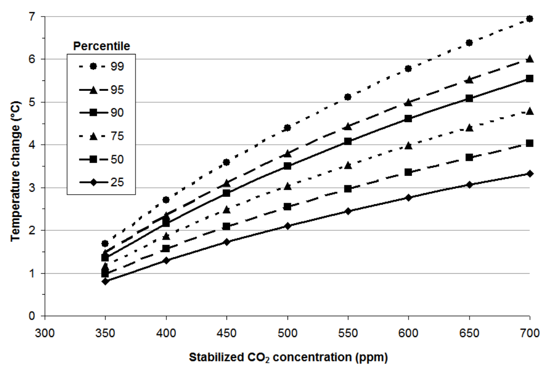
[15]Ultimately, stabilized CO_2 concentrations may not be the most helpful way of framing GHG goals (e.g., Allen et al., 2009; Matthews et al., 2009; Boykoff et al., 2010), but they do give an important sense of the risks implied by different types of emission pathways.



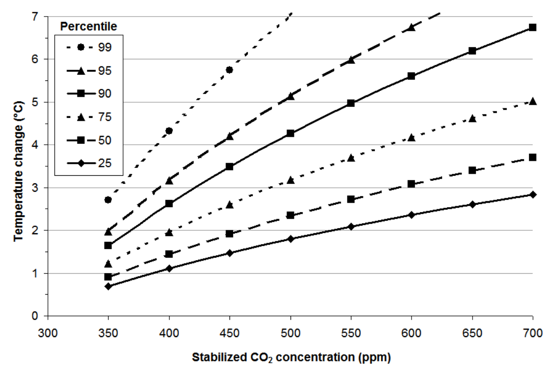
(a) From prior combination 3



(b) From prior combination 5



(c) Calculated from Annan and Hargreaves (2006)



(d) Calculated from Hegerl et al. (2006)

Figure 1.11: The posterior percentile temperature change relative to pre-industrial levels for different CO_2 concentrations. Percentile temperature change is calculated from Table 1.3 and equation (1.2). Prior combination 3 allows models to be incomplete and to share biases, and prior combination 5 does not. Also, data from Annan and Hargreaves (2006) and Hegerl et al. (2006), calculated using equation (1.2).

Table 1.3: The posterior percentile values for the aggregate feedback factor (Feedback-at-Risk, or FaR), and the aggregate feedback factor expected conditional on exceeding the posterior percentile value (Conditional Feedback-at-Risk, or CFaR).

	From Prior 3		From Prior 5	
Percentile	FaR	CFaR	FaR	CFaR
25	0.38	0.58	0.58	0.62
50	0.51	0.64	0.60	0.64
75	0.62	0.72	0.63	0.65
90	0.72	0.80	0.65	0.67
95	0.79	0.86	0.67	0.69
99	0.90	0.94	0.70	0.72

(slightly above present levels) and still have less than a 5% chance of exceeding the 2°C target. If, on the other hand, models are believed to possibly have shared biases and omissions (prior combination 3), then, before accounting for the effects of non-CO_2 GHGs or of aerosols, even stabilizing CO_2 concentrations at current levels leaves a 10% chance of exceeding the 2°C target. Policymakers' 2°C target may therefore support significant near-term abatement and eventual net negative emissions (Lemoine et al., 2011).

1.8 Conclusion

This paper elaborates an extensible method for combining studies of feedback factors to produce posterior distributions for aggregate feedbacks and temperature change in response to stabilized radiative forcing. Assumptions about shared model omissions and shared model biases are crucial for the positive tails of temperature change distributions and for the sensitivity of CO_2 concentration targets to temperature risk tolerance. Beyond obtaining observations with uncorrelated structural biases for use in constraining posterior distributions, further work could adopt more complex representations of model dependencies (Jun et al., 2008), could include total system constraints (Urban and Keller, 2009), could explore alternate types of prior beliefs, and could refine prior beliefs about shared structural biases and about unknown and unmodeled feedbacks. Further work could also develop temperature change distributions for planned emission pathways by including uncertainty about the operation of CO_2 sinks in response to changing CO_2 concentrations and uncertainty in monitoring negotiated emission allocations. A robust Bayesian approach may help to address the difficulty in choosing the "right" prior distribution for feedbacks and shared biases (e.g., Borsuk and Tomassini, 2005; Tomassini et al., 2007), and this paper's use of multiple priors could complement an ambiguity aversion framework for decision-making (e.g., Klibanoff et al., 2005, 2009). Finally, to make them more useful for adaptation work and impacts assessments, these probability distributions should be extended to consider transient climate change by including uncertainty about heat uptake by oceans, uncertainty about emission paths, and uncertainty about the timescales over which feedbacks operate (Hall, 2007; Knutti et al., 2008; Baker and Roe, 2009; Roe, 2009).

1.9 Appendix: Estimating the carbon cycle feedback factor using coupled climate-carbon cycle models

This appendix explains how to adapt the results of Friedlingstein et al. (2006) and Cadule et al. (2009) for use in the standard feedback model described in Roe (2009) and in section 3. It shows how their estimated feedback factors differ from f_{cc}, the feedback of interest, and explains how to combine their reported values with results from Soden et al. (2008) to estimate f_{cc}.

Let f_{cc} be the carbon cycle feedback and f_k be any of the other $K-1$ feedbacks. Define:

$$\alpha \equiv \lambda_0 \left.\frac{\partial R}{\partial \ln C}\right)_{x_{k,1\leq k\leq K-1}} \qquad \text{and} \qquad \eta \equiv \frac{\mathrm{d}\ln C_{T_c}}{\mathrm{d}T_c}$$

where x_k is the climate field for feedback process k. Let ΔC_u be the change in atmospheric CO_2 levels when the carbon cycle is uncoupled from a climate model (i.e., when the carbon cycle does not respond to changing temperatures), and let ΔC_T be the change in CO_2 levels resulting from temperature change ΔT. The total change in CO_2 levels in a coupled climate-carbon cycle model is therefore $\Delta C_c = \Delta C_u + \Delta C_{T_c}$, where ΔT_c is the temperature change in a coupled model. Because ΔR_f comes from the change in CO_2 levels as determined by an uncoupled carbon cycle model, we have exogenous forcing $\Delta R_f = \lambda_0^{-1}\alpha\Delta\ln C_u$. Then, from equation (1.1):

$$\Delta T_c = \alpha\Delta\ln C_u + \lambda_0\sum_{k=1}^{K-1}(c_k\Delta T_c) + \alpha\eta\Delta T_c = \frac{\alpha\Delta\ln C_u}{1-\sum_{k=1}^{K-1} f_k - \alpha\eta} \tag{1.9}$$

The carbon cycle feedback factor f_{cc}, the parameter of ultimate interest, is:

$$f_{cc} = \alpha\eta \tag{1.10}$$

Note that (1.10) corresponds to the equations in Friedlingstein et al. (2003, 2006) for their carbon cycle feedback factor g, except adjusting the definition of η to account for radiative forcing being proportional to log CO_2 levels.

Friedlingstein et al. (2003, 2006) calculated f_{cc} using estimates of α labeled herein as $\hat{\alpha}$. I show that $\hat{\alpha} \neq \alpha$ because they implicitly took the operation of other feedbacks in an uncoupled model to be part of the reference system. Friedlingstein et al. (2006) estimated $\hat{\alpha}$ from a regression using coupled model output via the equation $\Delta T_c = \hat{\alpha}\Delta C_c$. Friedlingstein et al. (2003) presented the regression as $\Delta T_c = \hat{\alpha}\Delta C_c + \Delta T_{ind}$, where ΔT_{ind} represents all non-CO_2 sources of temperature change and serves as a constant in the regression. These studies estimated carbon cycle feedbacks using regressions that attributed all of the variation in temperature to the variation in coupled model CO_2 concentrations. However, as seen in equation (1.9), changes in CO_2 concentrations actually also operate through other feedbacks to change temperature. As long as these other feedbacks are net positive, attributing to carbon cycle feedbacks the variation in CO_2 induced by these other feedbacks inflates $\hat{\alpha}$ and so, via equation (1.10), also inflates the estimate of the carbon cycle feedback factor. Imagine that temperature were proportional to changes in CO_2 concentrations, so replace $\Delta \ln C_u$ in equation (1.9) with ΔC_u and make an analogous replacement in the definition of η. Using these (incorrect) replacements and substituting for $\eta\Delta T_c = \Delta C_{T_c}$, for $\Delta C_u = \Delta C_c - \Delta C_{T_c}$, and for $\Delta T_c = \hat{\alpha}\Delta C_c$ in equation (1.9) gives α in terms of $\hat{\alpha}$:

$$\alpha = \hat{\alpha}\left(1 - \sum_{k=1}^{K-1} f_k\right) \tag{1.11}$$

Cadule et al. (2009) corrected Friedlingstein et al. (2006) by treating the change in temperature as proportional to the change in log CO_2 concentrations. They bypassed the use of α and instead calculated $\hat{f}_{cc}$, their estimate of f_{cc}, directly from changes in CO_2 concentrations in coupled and uncoupled model runs, but $\hat{f}_{cc} \neq f_{cc}$ due to the same reference system complication encountered by Friedlingstein et al. (2003, 2006). Let the coupled models' temperature change be given by:

$$\Delta T_c = \Delta T_u + \hat{f}_{cc}\Delta T_c \tag{1.12}$$

This yields the representation that Cadule et al. (2009) used:

$$\hat{f}_{cc} = \frac{\Delta T_c - \Delta T_u}{\Delta T_c} \tag{1.13}$$

However, equation (1.9) and the relation $\Delta T_u = \alpha\Delta \ln C_u + \lambda_0 \sum_{k=1}^{K-1} c_k \Delta T_u$ imply:

$$\Delta T_c = \Delta T_u + \sum_{k=1}^{K-1} f_k(\Delta T_c - \Delta T_u) + f_{cc}\Delta T_c \tag{1.14}$$

Substituting for $\Delta T_c - \Delta T_u$ from equation (1.12) and simplifying yields:

$$f_{cc} = (1 - \sum_{k=1}^{K-1} f_k)\hat{f}_{cc} \tag{1.15}$$

Table 1.4: The coupled models from Friedlingstein et al. (2006) and Cadule et al. (2009) with their corresponding uncoupled models from Soden et al. (2008).

Coupled model	Uncoupled model
CSM-1	NCAR CCSM3
FRCGC	MIROC MEDRES
HadCM3LC	UKMO HADCM3
IPSL-CM4-LOOP	IPSL
LLNL	NCAR PCM1
MPI	MPI ECHAM5
BERN-CC	–
CLIMBER	–
IPSL-CM2C	–
UMD	–
Uvic-2.7	–

Again, the given estimate of carbon cycle feedbacks includes the operation of the feedbacks in response to the increased temperature resulting from increased CO_2 levels. For our purposes, Cadule et al. (2009) thus overestimated the strength of carbon cycle feedbacks when other feedbacks are net positive, and knowing $\sum_{k=1}^{K-1} f_k$ enables identification of f_{cc} from their reported $\hat{f}_{cc}$.

The data points for f_{cc} reported in Table 1.2 and Figure 1.4 come from matching coupled models described in Friedlingstein et al. (2006) with uncoupled models described in Soden et al. (2008) (Table 1.4). Models match if they use closely related ocean-atmosphere general circulation models (OAGCMs). Some coupled models do not have a close match in the Soden et al. (2008) set of models because, for instance, some coupled models are not tied to an OAGCM but to a simpler model. For coupled models with matches, the sum of the albedo, cloud, and water vapor-lapse rate feedbacks in the uncoupled models enables use of equation (1.15) to adjust the carbon cycle feedback reported in Cadule et al. (2009).

Chapter 2

Paleoclimatic warming increased carbon dioxide concentrations[1]

If climate-carbon feedbacks are positive, then warming causes changes in carbon dioxide (CO_2) sources and sinks that increase CO_2 concentrations and create further warming. Previous work using paleoclimatic reconstructions has not disentangled the causal effect of interest from the effects of reverse causality and autocorrelation. The response of CO_2 to variations in orbital forcing over the past 800,000 years suggests that millennial-scale climate-carbon feedbacks are significantly positive and significantly greater than century-scale feedbacks. Feedbacks are also significantly greater on 100 year timescales than on 50 year timescales over the past 1,500 years. Posterior probability distributions implied by coupled models' predictions and by these paleoclimatic results give a mean of 0.03 for the non-dimensional climate-carbon feedback factor and a 90% chance of its being between -0.04 and 0.09. The 70% chance that climate-carbon feedbacks are positive implies that temperature change projections tend to underestimate an emission path's consequences if they do not allow the carbon cycle to respond to changing temperatures.

2.1 Introduction

Climate-carbon (or carbon cycle) feedbacks control how carbon dioxide (CO_2) concentrations respond to changing temperatures (Friedlingstein et al., 2006; Gregory et al., 2009). Positive feedbacks indicate that increased surface temperatures cause changes in CO_2 sources

[1]Originally published as:
Lemoine, D.M. 2010. Paleoclimatic warming increased carbon dioxide concentrations. Journal of Geophysical Research 115: D22122. doi:10.1029/2010JD014725

and sinks that in turn further increase surface temperatures (Cox et al., 2000; Heimann and Reichstein, 2008). Because other climate change feedbacks are thought to be positive on net (Bony et al., 2006; Soden et al., 2008), and because feedbacks add linearly but impact temperature nonlinearly (Torn and Harte, 2006; Roe and Baker, 2007; Roe, 2009), constraining the range of climate-carbon feedbacks is important for constraining temperature change projections and for climate risk assessments (Plattner et al., 2008; Huntingford et al., 2009). However, while models that couple the carbon cycle and the climate system can provide some insight into the possible magnitude of these feedbacks, the number and complexity of the interlinked processes restrict the amount of information that can be gleaned from models alone (Lemoine, 2010a).

Estimates from paleoclimatic data can provide an alternate source of information about the scale of feedbacks that may operate under anthropogenic warming. While differences in boundary conditions and in the type of forcing mean that paleoclimatic data are unlikely to correctly describe the earth system's response to ongoing anthropogenic greenhouse gas forcing, their biases in the anthropogenic application might be largely uncorrelated with those impacting coupled models' predictions (Lemoine, 2010a). Paleoclimatic estimates can therefore complement models' predictions in the construction of a probability distribution for climate-carbon feedbacks.

This paper estimates climate-carbon feedback strength over past ice age cycles and over the past two millennia. It uses changes in insolation due to orbital variations to identify the response of atmospheric CO_2 concentrations to changes in temperature over the previous 800,000 years. The results indicate that climate-carbon feedbacks were probably positive over past ice ages and over the past two millennia. The magnitude depends on the timescale of interest but, over millennial timescales, is comparable to coupled models' predictions of the carbon cycle's response to anthropogenic greenhouse gas forcing. The temperature change produced by a given emission path is therefore probably greater than suggested by climate sensitivity metrics that do not allow the carbon cycle to respond to changing temperatures.

2.2 Assessing feedback strength

The equilibrium temperature change ΔT due to a change in radiative forcing can be represented as:

$$\Delta T = \frac{\lambda_0 \Delta R_f}{1 - \sum_{k=1}^{K} c_k \lambda_0} = \frac{\lambda_0 \Delta R_f}{1 - F} \tag{2.1}$$

where λ_0 is the temperature change per unit of radiative forcing in the reference system upon which feedbacks operate, ΔR_f is the exogenous change in radiative forcing produced by increased GHG concentrations, and non-dimensional $f_k \equiv c_k \lambda_0$ gives the influence of feedback process k (Roe, 2009). This representation assumes that feedback processes are linear over the relevant temperature range and are defined so that they interact only through their effects on temperature. When positive, f_k may be interpreted as the fraction of total warming due to feedback process k.

Each feedback factor f_k can be decomposed into the product of the total change in climate field α_k due to a unit change in temperature and the change in radiative forcing due to a unit change in climate field α_k when other climate fields are held fixed (Roe, 2009). In the case of climate-carbon feedbacks f_{cc} affecting CO_2 concentrations, this gives:

$$f_{cc} \equiv f_k = \lambda_0 \left\{ \left. \frac{\partial R}{\partial \ln CO_2} \right)_{\alpha_{j,j \neq k}} \frac{d \ln CO_2}{dT} \right\} \tag{2.2}$$

where climate-carbon feedbacks are feedback process k. CO_2 concentrations are represented by their log because radiative forcing increases approximately linearly with the log of CO_2, yielding $\left. \frac{\partial R}{\partial \ln CO_2} \right)_{\alpha_{j,j \neq k}} = 5.35$ W m^{-2} (ln ppm)$^{-1}$ (Ramaswamy et al., 2001: Table 6.2). λ_0 is approximately 0.315 K (W m^{-2})$^{-1}$ (Soden et al., 2008). Estimating the climate-carbon feedback factor f_{cc} therefore primarily requires estimating $\psi \equiv d \ln CO_2 / dT$, or the effect of a unit of temperature change on CO_2 concentrations. Coupled climate-carbon cycle models have predicted this term (Friedlingstein et al., 2003, 2006; Cadule et al., 2009), but these models provide limited information because they only include a subset of known carbon cycle processes and are vulnerable to the possibility of shared model biases (Luo, 2007; Tebaldi and Knutti, 2007; Lemoine, 2010a).

Paleoclimatic estimates can provide an important additional source of information with biases largely independent of models' shared biases, but empirical estimation is complicated by the degree to which earth system components are intertwined, by the incompleteness of climatic records, and by the inability to run full-scale controlled experiments. Four studies have attempted to constrain climate-carbon feedbacks from temperature and CO_2 reconstructions. Scheffer et al. (2006) considered the last millennium's Little Ice Age (LIA), and Torn and Harte (2006) used the last 360,000 years as recorded by the Vostok ice core. Frank et al. (2010) estimated the response of CO_2 to temperature for three time periods in the past millennium. An ensemble of temperature and CO_2 reconstructions produced a frequency distribution for ψ. This distribution may be interpreted as a probability distribution for ψ if one assumes that the reconstructions properly sample the space of possible worlds. Finally, Cox and Jones (2008) constrained climate-carbon feedback strength by determining which values are consistent with the output of coupled climate-carbon cycle models run using twentieth century data, with the results of matching coupled models to observed interannual variability, and with a LIA analysis closely related to that of Scheffer et al. (2006).

Crucially, these four studies rely on univariate regressions of CO_2 on temperature that may contain biases from reverse causality and autocorrelation (Appendix A). A univariate regression cannot disentangle whether high CO_2 levels accompany high temperatures because higher CO_2 causes higher temperatures, because higher temperatures cause higher CO_2, or because they are each being driven by, for instance, previous periods' CO_2 and temperature.

Because feedback estimation is concerned with the response of CO_2 to an exogenous increase in temperature, it is important that paleoclimatic studies isolate the response of CO_2 to temperature from the more general correlation estimated by a univariate regression. The present study seeks to isolate the causal effect of temperature on CO_2 by looking at the response of CO_2 to variations in temperature that were unlikely to be caused by variations in CO_2.

2.3 Methods: Estimated equations

The present study estimates climate-carbon feedbacks over four timescales: millennia, centuries, 100 y, and 50 y. It seeks to generate estimates that are free of simultaneous equations (or reverse causality) bias and omitted variables bias. First, it aims to avoid simultaneous equations bias by using orbital forcing as an instrument for temperature over the longer timescales (Appendix A). A good instrument is correlated with temperature but only affects the coeval CO_2 concentration through its effect on temperature. In other words, using this instrument isolates a "good" portion of the variation in temperature—a portion that is believed not to be caused by changes in CO_2—and ignores the rest. A good instrument avoids the problem of imputing the causal effect of temperature on CO_2 from data that actually reflects the greenhouse effect of CO_2 on temperature.

The key hypotheses for the validity of an orbital forcing instrument are that: a) changes in orbital forcing cause changes in temperature, but b) do not affect CO_2 levels except through their effect on temperature. If these hypotheses hold, then we can replace the actual temperature record with one predicted from orbital forcing data and believe that any remaining correlation with the CO_2 record is due to the effect of temperature on CO_2. The first hypothesis is supported by the Milankovitch theory of glacial cycles, according to which summer insolation in the northern hemisphere's high latitudes controls both hemispheres' temperature on millennial timescales (Milanković, 1941; Hays et al., 1976; Berger, 1992). Variations in summer insolation might have this effect because nonlinearities in the climate system can amplify the direct effect on ice sheets and snow accumulation. Importantly for the choice of which insolation time series to use, some have instead argued that the true trigger for deglaciation is the timing of spring insolation in the northern hemisphere (Hansen et al., 2007) or that Antarctic temperatures are more tightly controlled by the duration of the local (southern hemisphere) summer (Huybers and Denton, 2008). While the hypotheses are difficult to distinguish empirically (Huybers, 2009) and the true mechanism may be more complex (Wolff et al., 2009), recent evidence does support a northern hemisphere trigger for Antarctic temperatures (Kawamura et al., 2007; Cheng et al., 2009). Further, several recent studies (Petit et al., 1999; Jouzel et al., 2007; Kawamura et al., 2007) used high latitude summer solstice insolation in the northern hemisphere as an indicator of orbital forcing, and ice core chronologies sometimes assume a linear response of climate to orbital forcing, whether defined via mid-June insolation at northern high latitudes (Parrenin et al., 2004) or via anticipated periodicity (Salamatin, 2000). Therefore, given that orbital forcing should affect temperature, the key condition becomes the hypothesis that it does not directly affect the CO_2 concentration. Because orbital forcing affects insolation the most at the poles and the least at the equator, and because the primary effect at the poles is on snow and ice melt (via temperature), orbital forcing's effect on the timing and spatial distribution of insolation may not be directly critical for important carbon sources and sinks. Variations in orbital forcing may in fact cause variations in temperature without affecting CO_2 concentrations except through these variations in temperature.

The second source of bias that univariate regressions are exposed to is omitted variables bias produced by correlation of time t temperature and CO_2 with previous temperature and CO_2. If not accounted for, such correlation with past climate states could induce correlation between time t temperature and CO_2 that univariate regressions include in their coefficient estimates. However, this correlation through previous climate states may not be the effect of interest in a feedback application. The present study seeks to minimize omitted variables bias by including lagged covariates in the regression. The estimated model assumes that temperatures and concentrations at times earlier than those included as covariates only affect the temperature and concentration at time t through their effect on the included covariates.

The present study does not eliminate a final source of bias. Measurement error in temperature data may be due to errors in measurement of isotopes, in inferences about local temperature from isotopes, in inferences about global temperature from local temperature, and in the assignment of relative dates to the recorded temperature and CO_2. This measurement error tends to push coefficient estimates towards zero (Appendix A). Further, gas diffusion processes mean that each CO_2 observation actually has a distribution of ages and an effective resolution of a few centuries (Spahni et al., 2003), which tends to reduce the variation useful for regression-based estimates. The remaining errors should therefore tend to bias the results towards finding no effect of temperature on CO_2.

The orbital forcing specification estimates the following equation:

$$C_t = \beta_0 + \sum_{i=0}^{2} \beta_{i+1} T_{t-i} + \beta_4 C_{t-1} + \epsilon_t \tag{2.3}$$

where C_t is the log of the CO_2 concentration at time t, T_t is the temperature at time t, and t is in thousands of years. C_{t-2} is not included as a covariate because CO_2 concentrations from 2000 years ago should only affect contemporary CO_2 concentrations via their effect on CO_2 concentrations and temperature 1000 years ago. Orbital forcing (O_t) in W m^{-2} instruments for T_t via the following first-stage regression:

$$T_t = \gamma_0 + \gamma_1 O_t + \sum_{i=1}^{2} \gamma_{i+1} T_{t-i} + \gamma_5 C_{t-1} + \nu_t \tag{2.4}$$

O_t and T_t have a correlation coefficent of 0.18, so, as required for valid use as an instrument, variation in orbital forcing is connected to variation in temperature. The estimated covariance matrix uses the Huber-White estimator that is robust to arbitrary heteroskedasticity. Importantly for the applicability of the statistical methods used here, the time series appear to be stationary (augmented Dickey-Fuller tests reject the unit root hypothesis at the $\alpha = 0.05$ level), which means that the mean and covariance are not changing over time. It is also important that the error term ϵ_t not be serially correlated, because serial correlation may mean that ϵ_t is correlated with C_{t-1} via its correlation with ϵ_{t-1}, which would violate the assumption of exogeneity of the covariates. We test for such serial correlation in the instrumental variable estimate by using a Cumby-Huizinga test, which fails to reject the null hypothesis of no serial correlation at the $\alpha = 0.20$ level. We therefore assume that ϵ_t is not serially correlated and that C_{t-1} is in fact exogenous for ϵ_t.

The resulting coefficients and covariance matrix enable estimation of feedbacks over two timescales. The feedback factor over a timescale of j time units is calculated from equation (2.2) using:

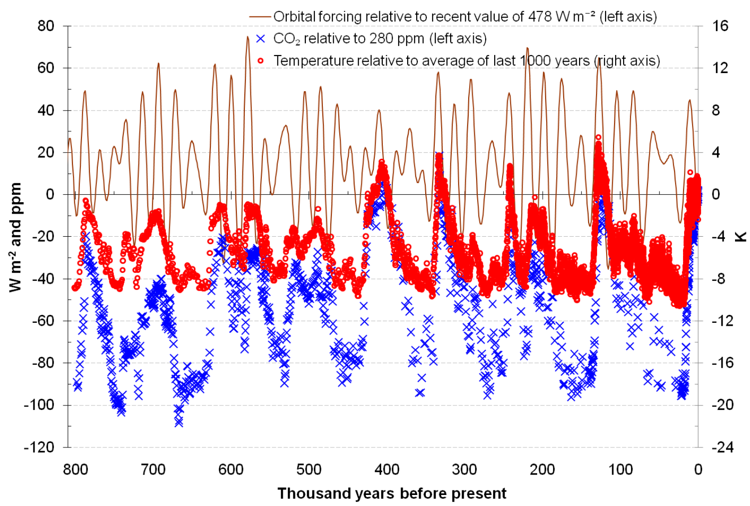
$$\psi \equiv \psi_j = \sum_{i=0}^{j} T_{t-i} + \sum_{k=1}^{j} \psi_{j-k} C_{t-k} \tag{2.5}$$

where C and T variables represent their estimated coefficients and $j \geq 0$. ψ_j is defined recursively, and ψ_0 is the coefficient on T_t. Thus, β_1 gives the effect of T_t on C_t, which is here labeled the century-scale response, and $\beta_1 + \beta_2 + \beta_4\beta_1$ gives the effect on C_t of an increase in temperature at time $t-1$ that is maintained at time t, which is here labeled the millennial response. Variance and covariance calculations use first-order linear approximations for the $\psi_{j-k}C_{t-k}$ terms.

The data are an 800 ky temperature record from the Antarctic EPICA Dome C core with the EDC3 age scale (Jouzel et al., 2007), an 800 ky composite CO_2 record drawn from that and other cores (Lüthi et al., 2008), and the calculations of Berger (1978) for orbital forcing at 60°N (Figure 2.1a). The similarity of this temperature record to those of the Vostok and Dome F cores implies that it may be indicative of general conditions over eastern Antarctica (Jouzel et al., 2007), and models suggest that Antarctic temperatures may track global temperatures (Masson-Delmotte et al., 2010). Figure 2.1b shows how including lagged variables as covariates alters the temperature-CO_2 relationship and how the instrument isolates a portion of the variation in T_t.

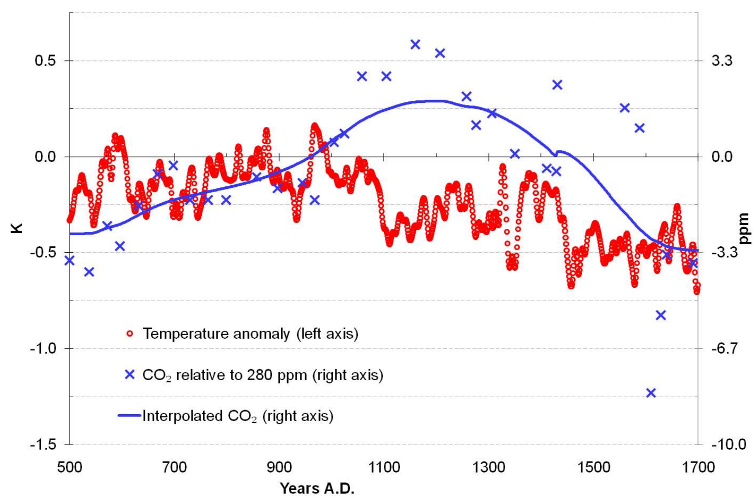
Estimating coefficients in several model specifications assesses the results' robustness to some types of specification error. In the base case and summer insolation specifications, the temperature and CO_2 data used are the observations closest to the endpoint of each 1000 year interval, while the averaged data specification uses the average of the previous 1000 years' observations. In the base case and averaged data specifications, the orbital forcing instrument is insolation in mid-June, but the summer insolation specification sums the insolation over June, July, and August.

The orbital forcing regressions estimate feedback strength over timescales of centuries or millennia, but it is also of interest to nearer-term climate projections to estimate climate-carbon feedbacks over shorter timescales. This requires a denser dataset than is available



[h]

(a)



(b)

Figure 2.1: (a) Mid-June orbital forcing at 60°N (Berger, 1978) instruments for the 800 ky EPICA Dome C temperature record (Jouzel et al., 2007) in a regression with data from a composite CO_2 record (Lüthi et al., 2008). (b) The 1500 y composite global temperature reconstruction (Mann et al., 2008) (EIV with HadCRUT3v) is used in a regression with interpolated CO_2 data from the Law Dome ice core (MacFarling Meure et al., 2006). All datasets are truncated at 1700 A.D. to avoid the Industrial Revolution.

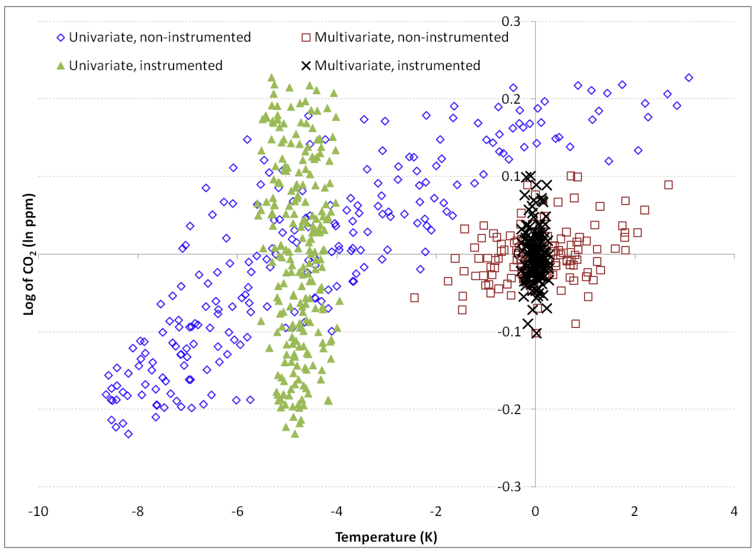


Figure 2.2: The relationship between temperature at time t (T_t) and the log of CO_2 at time t (C_t) estimated by instrumented and non-instrumented univariate and multivariate regressions over the 800 ky paleoclimatic reconstructions. The non-instrumented univariate regression shows demeaned C_t against T_t. The non-instrumented multivariate regression shows the residuals from a regression of C_t on the covariates excluding T_t against the residuals from a regression of T_t on the covariates. The instrumented regressions are similar except replacing T_t with its predicted value from the appropriate first-stage regression.

from ice cores. We therefore use composite temperature records from 500 A.D. through 1700 A.D. (Mann et al., 2008) and the CO_2 record from the Law Dome ice core (MacFarling Meure et al., 2006). The CO_2 record is made denser by first using Friedman's supersmoother algorithm under the assumption that CO_2 concentrations only change slowly and smoothly over century-scale timespans prior to 1700 A.D. Shape-preserving piecewise cubic interpolation then fills in values for missing years (Figure 2.1b). With t on the order of decades rather than in thousands of years, it is important to include several lagged terms because more distant lags may now affect time t variables directly (e.g., Schimel et al., 1996). The estimated model for shorter-term feedbacks is:

$$\Delta C_t = \sum_{i=0}^{k} \beta_{1+i} \Delta T_{t-i} + \sum_{j=1}^{k} \beta_{1+k+j} \Delta C_{t-j} + \epsilon_t \tag{2.6}$$

where Δ indicates a first difference (so $\Delta C_t = C_t - C_{t-1}$) and where $k = 11$ when the timestep for t is 10 years while $k = 5$ when the timestep for t is 25 years. Differencing the data makes it stationary (augmented Dickey-Fuller tests reject the unit root hypothesis at the $\alpha = 0.10$ level), and Durbin's alternative test—a standard test for serial correlation in Ordinary Least Squares estimates—fails to reject the null hypothesis of no serial correlation at the $\alpha = 0.50$ level. We calculate the effect of a 50-year and 100-year maintained increase in temperature from equation (2.5) using the estimated coefficients and heteroskedasticity-robust covariance matrix.

These subcentury timescale specifications do not instrument for T_t for two reasons. First, simultaneous equations bias should be small. This is because any unobserved sources of variation in CO_2 levels that appear between time $t-1$ and time t should be small and may not have enough time to fully affect T_t. Second, despite significant first-stage coefficients, weak instrument tests indicate potential problems with the use of solar activity from Steinhilber et al. (2009) and Delaygue and Bard (2010) as an instrument for T_t. Even if simultaneous equations bias is nonzero, it is probably sufficiently small that the Ordinary Least Squares estimate is preferable to estimation with a weak instrument.

2.4 Results

The orbital forcing specifications indicate that expected millennial-scale climate-carbon feedbacks are probably positive ($p<0.001$), acting to amplify anthropogenic warming (Table 2.1). Their 95% confidence intervals are in the range of 0.02 to 0.05 (Figure 2.3), which is comparable to the predictions of the coupled climate-carbon cycle models described in Friedlingstein et al. (2006). However, in line with the anticipated effects of biases introduced to previous work by reverse causality and autocorrelation, this range is on the low end of previous paleoclimatic estimates. Climate-carbon feedbacks are statistically greater over millennial timescales than over timescales of centuries ($p<0.001$), and for either 10-year or 25-year timesteps, climate-carbon feedbacks are statistically greater over 100 year

Table 2.1: Estimation results for the non-dimensional climate-carbon feedback factor f_{cc}.

Timescale	Specification	Data's timestep	n [a]	f_{cc}	s.e. [b]	p [c]
Millennia	Base case	1000 y	525	0.03	(0.009)	0.0001
Millennia	Summer insolation	1000 y	525	0.03	(0.009)	0.0007
Millennia	Averaged data	1000 y	536	0.03	(0.01)	0.0001
Centuries	Base case	1000 y	525	0.009	(0.01)	0.4
Centuries	Summer insolation	1000 y	525	0.006	(0.01)	0.6
Centuries	Averaged data	1000 y	536	0.002	(0.02)	0.9
100 y	–	10 y	109	0.02	(0.005)	0.003
100 y	–	25 y	43	0.01	(0.009)	0.1
50 y	–	10 y	109	0.005	(0.002)	0.006
50 y	–	25 y	43	0.005	(0.004)	0.2

[a] Number of observations
[b] Standard errors are robust to arbitrary heteroskedasticity.
[c] Two-tailed p-value for the null hypothesis that f_{cc} is equal to 0

timescales than over 50 year timescales ($p<0.001$). Each first-stage regression produces a coefficient on the orbital forcing instrument that is significantly different from 0 ($p<0.001$), and heteroskedasticity-robust Kleibergen-Paap F statistics greater than 15 confirm that the orbital forcing instrument should not pose weak instrument problems.

Most coefficient estimates are fairly stable across orbital forcing specifications and have the expected signs, indicating that the general model is robust to the specifications considered here (Table 2.2). Both millennial and century-scale feedback estimates are also relatively stable over different 200 ky sections of the datasets, with the main variations correlated with variations in the strength of the instrument (Figure 2.4). The paper's main findings therefore should not be highly sensitive to the choice of time period.

Univariate regressions and a non-instrumented multivariate regression help assess the possible importance of omitted variables bias and simultaneous equations bias (Table 2.3). Failing to disentangle the (positive) causal effect of CO_2 on temperature should make the effect of temperature on CO_2 seem stronger and reduce uncertainty about its point estimate. Indeed, as expected, the non-instrumented regressions produce greater feedback estimates with smaller standard errors. While the instrumented univariate regression does produce a similar point estimate and standard error for the coefficient on T_t as do the instrumented multivariate regressions, it is less useful for estimating millennial feedbacks because it does not allow previous temperature or CO_2 concentrations to affect time t values.

In estimation of decadal-scale feedbacks, coefficients on the more recent CO_2 levels are

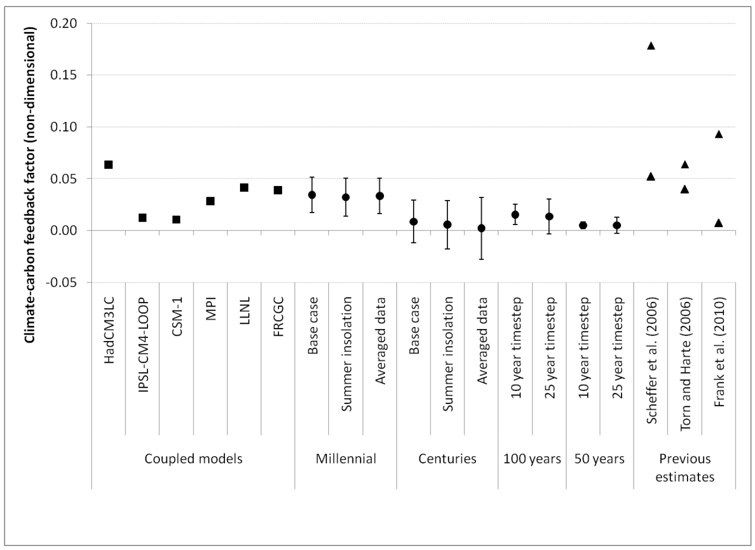


Figure 2.3: Estimates of the climate-carbon feedback factor f_{cc}. Coupled climate-carbon cycle models are as described in Friedlingstein et al. (2006), and their plotted points are the average of the results from Lemoine (2010a) for the three radiative kernels. Error bars show the 95% confidence intervals for this paper's paleoclimatic estimates. Previous paleoclimatic estimates are converted to feedback form using the factor of 1.2 K (275 ppm)$^{-1}$ from Torn and Harte (2006) and, in the case of Frank et al. (2010), indicate the range of "likely" values. These previous paleoclimatic estimates assumed that radiative forcing increases linearly with CO_2 rather than with the log of CO_2.

Table 2.2: Coefficient estimates and standard errors from the orbital forcing specifications.[a]

Specification	n[b]	Second stage T_t [c]	T_{t-1} [c]	T_{t-2} [c]	C_{t-1}	$Const$	First stage O_t [d]	T_{t-1}	T_{t-2}	C_{t-1} [e]
Base case	525	0.005	0.01	-0.01***	0.9***	0.8***	0.007***	1***	-0.2***	2**
		(0.006)	(0.007)	(0.002)	(0.03)	(0.1)	(0.002)	(0.06)	(0.06)	(0.8)
Summer insolation	525	0.003	0.01	-0.01***	0.9***	0.8***	0.00009***	1***	-0.2***	2**
		(0.007)	(0.008)	(0.002)	(0.03)	(0.2)	(0.00002)	(0.06)	(0.06)	(0.8)
Averaged data	536	0.001	0.02	-0.01***	0.9***	0.6***	0.005***	1***	-0.4***	0.6
		(0.009)	(0.01)	(0.004)	(0.02)	(0.1)	(0.001)	(0.05)	(0.05)	(0.6)

[a] Standard errors (in parentheses) are robust to arbitrary heteroskedasticity. Two-tailed p-values are for the null hypothesis that the true coefficient is equal to 0: * means $p < 0.1$, ** means $p < 0.05$, and *** means $p < 0.01$. t is in 1000 years.
[b] Number of observations
[c] Units of (ln ppm CO_2) K^{-1}
[d] Units of K $(W\ m^{-2})^{-1}$
[e] Units of K (ln ppm $CO_2)^{-1}$

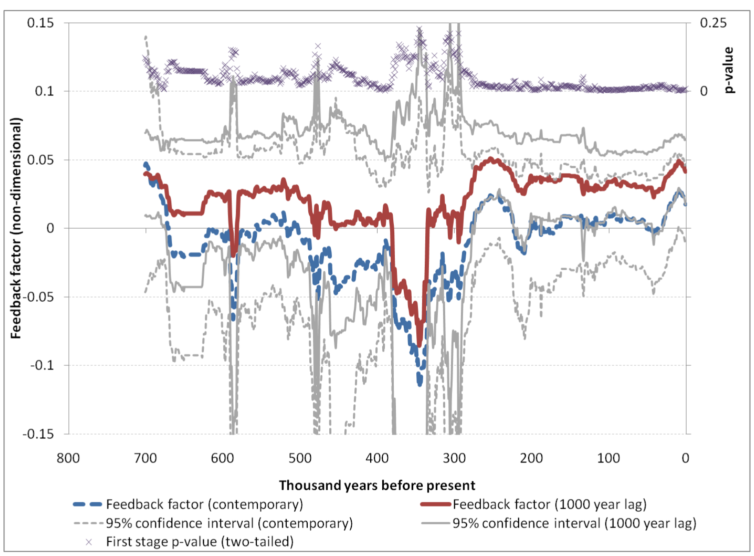


Figure 2.4: Estimates of century-scale and millennial climate-carbon feedbacks f_{cc} are relatively stable over each 200 ky window in the dataset for which the instrument's strength is stable (as indicated by the two-tailed p-value on the coefficient of the orbital forcing instrument).

Table 2.3: Coefficient estimates and standard errors in versions of the base case orbital forcing specification without using instruments and/or without including lagged variables as covariates.[a]

	n[b]	T_t[c]	T_{t-1}[c]	T_{t-2}[c]	C_{t-1}	$Const$	O_t[d]	f_{cc} Millennial[e]	f_{cc} Centuries
Univariate, non-instrumented	638	0.03***	–	–	–	6***	–	0.1***	0.06***
		(0.0002)	–	–	–	(0.008)	–	(0.005)	(0.003)
Univariate, instrumented	638	0.007	–	–	–	5***	0.02**	0.02	0.01
		(0.01)	–	–	–	(0.06)	(0.01)	(0.02)	(0.04)
Multivariate, non-instrumented	525	0.02***	-0.0006	-0.009***	0.8***	0.9***	–	0.05***	0.03***
		(0.002)	(0.002)	(0.002)	(0.03)	(0.1)	–	(0.003)	(0.003)

[a] Standard errors (in parentheses) are robust to arbitrary heteroskedasticity in the multivariate case and also to arbitrary autocorrelation in the univariate cases. Two-tailed p-values are for the null hypothesis that the true coefficient is equal to 0: * means $p < 0.1$, ** means $p < 0.05$, and *** means $p < 0.01$. t is in 1000 years. The instrumented univariate case has a robust Kleibergen-Papp F statistic of 6, indicating the potential for a weak instrument problem.
[b] Number of observations
[c] Units of (ln ppm CO_2) K^{-1}
[d] First-stage regression result with units of K $(W\ m^{-2})^{-1}$
[e] In the univariate cases, assumes that the coefficient on T_{t-1} is certainly equal to zero.

often significant while the other coefficients are usually not significant (Table 2.4). This accords with the intuition that, over such short timescales and with the correspondingly small variation in CO_2 and temperature over each timestep, the time t CO_2 level should be almost wholly determined by the previous period's CO_2 level. Mann et al. (2008) provided several composite temperature records calculated using different instrumental records and combined using different statistical techniques. All results reported in this paper use the reconstruction resulting from their error-in-variables estimation procedure and calibrated using HadCRUT3v instrumental land and ocean hemispheric means. Using the other error-in-variables temperature reconstruction from Mann et al. (2008) does not substantially affect the results, but using the reconstruction developed using the composite plus scale methodology tends to produce estimates that are not significantly different from 0.

2.5 Discussion

The point estimates and standard errors provide information about the sampling distribution of the mean, but the probability distribution for the feedback factor is more important. Appendix B describes how to develop a probability distribution by extending the hierarchical Bayes framework of Lemoine (2010a) to combine this paper's base case empirical estimates with coupled models' predictions. The posterior distribution implied by the empirical studies is similar to the one implied by coupled models' output, but considering both types of data together can further constrain the posterior distribution (Figure 2.5). With only data from coupled models, it is difficult to disentangle the true feedback factor from the biases shared among those models, but empirical estimates provide information about the true feedback factor that is affected by a different set of biases. The posterior distribution resulting from using both types of data has a mean of 0.03 and 5th and 95th percentile values of -0.04 and 0.09. It also indicates a roughly 70% chance that climate-carbon feedbacks are positive, thereby reinforcing other feedbacks such as those due to changes in albedo and water vapor content. Instead of obtaining point estimates and standard erorrs, future work could develop probability distributions directly from paleoclimatic data and then combine those with coupled models' predictions.

The proper application of this paper's empirical feedback estimates to anthropogenic climate change depends on the question of interest. Feedback strength may vary with timescale, and future feedbacks will operate in a world with different boundary conditions and with radiative forcing changing with a scale and speed not represented in paleoclimatic data or in data used to tune coupled models. Further, feedback strength may depend on the pace of climate change, and uncertainty about concentration-carbon feedbacks may be more important to the total carbon cycle response than is uncertainty about climate-carbon feedbacks (Gregory et al., 2009). A complete accounting of carbon cycle uncertainty must include these factors as well as concerns about irreversible changes.

Paleoclimatic records suggest that climate-carbon feedbacks are positive, despite the

Table 2.4: Coefficient estimates and standard errors from the specifications used to estimate decadal-scale feedbacks.[a]

	10 y timestep (n=109)		100 y timestep (n=43)	
Parameter	Estimate	S.E.	Estimate	S.E.
ΔT_t	0.0001**	(0.00005)	0.0003	(0.0004)
ΔT_{t-1}	0.00008	(0.00006)	0.0008	(0.0005)
ΔT_{t-2}	0.0002**	(0.00009)	0.0006	(0.0005)
ΔT_{t-3}	0.0001*	(0.00007)	0.001**	(0.0005)
ΔT_{t-4}	0.00004	(0.00007)	-0.0001	(0.0004)
ΔT_{t-5}	0.0002	(0.0001)	0.0005	(0.0003)
ΔT_{t-6}	0.0002*	(0.00009)		
ΔT_{t-7}	0.0002	(0.0001)		
ΔT_{t-8}	0.0002*	(0.00008)		
ΔT_{t-9}	-0.0002**	(0.0001)		
ΔT_{t-10}	0.00007	(0.0001)		
ΔT_{t-11}	0.00004	(0.00006)		
ΔC_{t-1}	1***	(0.1)	1***	(0.3)
ΔC_{t-2}	-0.7***	(0.2)	-0.2	(0.4)
ΔC_{t-3}	0.7**	(0.3)	0.2	(0.3)
ΔC_{t-4}	-0.6**	(0.2)	-0.2	(0.1)
ΔC_{t-5}	0.4	(0.3)	-0.01	(0.09)
ΔC_{t-6}	-0.3	(0.3)		
ΔC_{t-7}	0.3	(0.3)		
ΔC_{t-8}	-0.2	(0.2)		
ΔC_{t-9}	0.2	(0.2)		
ΔC_{t-10}	-0.2*	(0.1)		
ΔC_{t-11}	0.06	(0.08)		

[a] Standard errors are robust to arbitrary heteroskedasticity. Two-tailed p-values are for the null hypothesis that the true coefficient is equal to 0: * means $p < 0.1$, ** means $p < 0.05$, and *** means $p < 0.01$. Coefficients on temperature terms are in units of (ln ppm CO_2) K^{-1}.

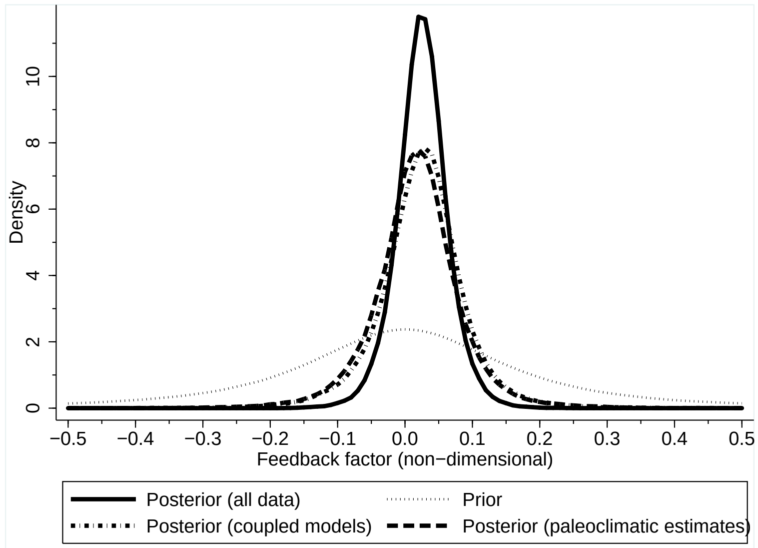


Figure 2.5: The posterior distributions for f_{cc} produced by the statistical framework from Lemoine (2010a) when the prior distributions are updated with output from coupled climate-carbon cycle models, with this paper's base case paleoclimatic estimates for the orbital forcing specifications, and with both the coupled models' output and this paper's base case paleoclimatic estimates.

presence of measurement error that should lead to underestimation of feedback strength. Obtaining more precisely dated paleoclimatic records with denser data could be crucial for better identification of feedback strength, and longer Holocene time series with denser data are important for estimation on subcentury timescales. It appears as if coupled models' feedback predictions are more apt than are the higher estimates of previous paleoclimatic work. Importantly, combining coupled models' output with this paper's empirical estimates sufficiently constrains climate-carbon feedbacks so that they might not be a dominant source of uncertainty about future temperature change. Temperature risk assessments are probably more dominated by the possibility of tipping points and of shared biases among models (O'Neill and Oppenheimer, 2004; Lenton et al., 2008; Lemoine, 2010a). However, climate policy analyses can be especially sensitive to the positive tail of temperature change distributions because damages may increase nonlinearly with the temperature index and because climate decision-makers are usually modeled as risk averse (Newbold and Daigneault, 2009; Weitzman, 2009). Because positive climate-carbon feedbacks thicken these policy-relevant positive tails, considering their existence and associated uncertainty is important not just for climate projections but also for economic assessments that may otherwise underestimate climatic risks.

2.6 Appendix: Sources of bias in estimating climate-carbon feedbacks

Previous empirical work estimated climate-carbon feedbacks using Little Ice Age data and Vostok ice core data (Scheffer et al., 2006; Torn and Harte, 2006; Cox and Jones, 2008; Frank et al., 2010). These studies ran univariate Ordinary Least Squares (OLS) regressions of CO_2 on temperature, but the estimates produced by such a regression are vulnerable to several sources of bias that complicate attempts to apply the results to the current global radiative forcing experiment. Adjusting them to use log concentrations, those univariate regressions may be represented as:

$$C_t = \mu + \beta T_t + \epsilon_t \tag{A}$$

where C_t is the log of the CO_2 concentration at time t, T_t is the temperature at time t, μ is a constant term, and ϵ_t is the random unobserved error at time t. The parameter of interest is β, which ideally gives $\partial C/\partial T$ or even dC/dT. The linearized full system may look more like:

$$\begin{cases} C_t &= \mu_C + \sum_{i=0}^{m} \beta_i T_{t-i} + \sum_{j=1}^{n} \gamma_j C_{t-j} + \eta_t \\ T_t &= \mu_T + \sum_{i=0}^{p} \alpha_i C_{t-i} + \sum_{j=1}^{q} \phi_j T_{t-j} + \nu_t \end{cases} \tag{B}$$

In this representation, CO_2 concentrations and temperature each depend on their own past values, on the past values of the other variable, on the constants μ_C and μ_T, and on the random errors η_t and ν_t. Here, the parameter of interest depends on the allowed time for carbon cycle responses, but it is either β_0 or some combination of the β, γ, α, and ϕ parameters that gives the effect of a maintained unit change in temperature on future log CO_2 concentrations.

Assume for the rest of this section that the parameter of interest is β_0, which may be the case if the data's timestep is larger than the timescale of interest in the feedback application. When the true system is (B), estimating β_0 via the univariate regression in equation (A) introduces three sources of bias via the correlation between ϵ_t and T_t. First, assume that the true system has, $\forall i > 0$, $\alpha_i = \beta_i = 0$ and, $\forall i$, $\gamma_i = \phi_i = 0$. In this case, previous CO_2 concentrations and previous temperatures would not affect current CO_2 concentrations and temperatures, but the current CO_2 concentration and the current temperature would affect each other. The simplified system of equations becomes:

$$\begin{cases} C_t &= \mu_C + \beta_0 T_t + \eta_t \\ T_t &= \mu_T + \alpha_0 C_t + \nu_t \end{cases} \tag{C1}$$

where $\eta_t = \epsilon_t$ from equation (A). Let b_0 be the OLS estimate of β_0 from equation (A) so that $\operatorname{plim} b = \beta_0 + \frac{Cov(T_t,\epsilon_t)}{Var(T_T)}$. If T_t is exogenous for C_t, then $Cov(T_t, \epsilon_t) = 0$ and b is a consistent estimator of β_0. However, from (C1), $Cov(T_t, \epsilon_t) = Cov(T_t, \eta_t) = \frac{\alpha_0}{1-\beta_0\alpha_0} Var(\epsilon_t)$. Because we know $\alpha_0 > 0$ (indeed, this is the greenhouse effect in this specification), the OLS estimate b is asymptotically biased upwards as long as ϵ_t is uncertain. Unobserved non-temperature factors that affect CO_2 levels through ϵ_t also affect temperature via the usual radiative forcing mechanism, which biases the OLS estimate of the effect of temperature on CO_2 by amplifying the relationship between observed temperature and observed CO_2. Measurement error in the CO_2 data is also subsumed in ϵ_t and thus can also produce simultaneous equations bias. This bias may be nonexistent if temperature is deemed not to respond to CO_2 on the timescale of interest (as Frank et al. (2010) and Scheffer et al. (2006) argued for the Little Ice Age) or if there is both no non-temperature driver of CO_2 and no measurement error for CO_2. Instrumental variables methods potentially enable one to avoid simultaneous equations bias without making such strong assumptions.

Second, replace the previous paragraph's assumptions with the assumption that, $\forall i$, $\alpha_i = 0$. This means that CO_2 does not affect temperature in the data of interest, which is an explicit reason Scheffer et al. (2006) and Frank et al. (2010) chose to study the Little Ice Age. In addition, assume that $\exists j > 0$ such that $\phi_j \neq 0$. The system of equations now becomes:

$$\begin{cases} C_t &= \mu_C + \sum_{i=0}^{m} \beta_i T_{t-i} + \sum_{j=1}^{n} \gamma_j C_{t-j} + \eta_t \\ T_t &= \mu_T + \sum_{j=1}^{q} \phi_j T_{t-j} + \nu_t \end{cases} \tag{C2}$$

Simultaneous equations bias does not appear if estimating β_0 in (C2) from equation (A), but the lagged variables create a different problem. In (A), the error term ϵ_t is a function of lagged temperature values when the true system is (C2). However, because previous temperatures affect the temperature observed at time t, $Cov(T_t, T_{t-i}) \neq 0$ for some $i > 0$, and because previous temperatures also affect CO_2 at time t but are omitted from the estimated system (A), we have $Cov(T_t, \epsilon_t) \neq 0$. The lagged temperatures act as omitted variables that bias estimates of β_0 in equation (A). Because these omitted variables are probably positively correlated with T_t and probably have positive coefficients in (C2), this bias probably also inflates positive estimates of β_0.

Third, replace the previous paragraphs' assumptions with the assumption that the true system has, $\forall i > 0$, $\beta_i = 0$ and, $\forall i$, $\alpha_i = \gamma_i = \phi_i = 0$. The true system becomes:

$$\begin{cases} C_t &= \mu_C + \beta_0 T_t + \eta_t \\ T_t &= \mu_T + \nu_t \end{cases} \tag{C3}$$

where η_t is uncorrelated with any time's temperature or with any previous CO_2 level. OLS estimation of β_0 via equation (A) would be consistent and unbiased with system (C3) if temperature were measured without error. However, temperature is actually measured but imperfectly. Let the observed temperature values be T_t^*, where:

$$T_t^* = T_t + w_t \tag{2.7}$$

w_t is a random variable that produces measurement error. Substituting into (C3), we get:

$$C_t^* = \mu_C + \beta_0 T_t^* + \eta_t' \tag{2.8}$$

$$\text{where} \qquad \eta_t' = \eta_t - \beta_0 w_t$$

Measurement error w_t in T_t induces nonzero correlation between η_t' and the observed T_t^*. If w_t has variance σ_w^2, we have:

$$Cov(T_t^*, \eta_t') = Cov(T_t + w_t, \eta_t - \beta_0 w_t) = -\beta_0 \sigma_w^2 \tag{2.9}$$

The random, unobserved measurement error in the temperature record biases the OLS estimate of β_0 towards zero ("attenuation bias"). This measurement error may be due to errors in measurement of isotopes, in inferences about local temperature from isotopes, in inferences about global temperature from local temperature, and in the assignment of relative dates to the recorded temperature and CO_2. Measurement error should be the primary source of bias remaining in the present study, and it is to some extent inescapable in work using data from limited paleoclimatic datasets.

2.7 Appendix: Hierachical Bayes model for combining coupled models' output with empirical estimates

This appendix outlines a statistical model which largely follows that described in Lemoine (2010a) but is adjusted to include a second group of studies (this paper's base case paleoclimatic estimates) that may have their own shared biases. Let f_{cc} represent the true value of the climate-carbon feedback factor and let θ_j represent the biases shared by group j (where j is an index indicating that studies are coupled models or paleoclimatic estimates). Crucially, assume that θ_1 and θ_2 are independent of each other, meaning that empirical studies' shared biases are assumed to be independent of those impacting coupled climate-carbon cycle models.

The empirical studies used here are the base case estimate of millennial climate-carbon feedbacks and the base case estimate of century-scale climate-carbon feedbacks. For each empirical study i, λ_{ij} represents the divergence between the object of the estimation procedure ($\hat{z}_i$) and the feedback of interest for projecting future temperature change (f_{cc}). λ_{ij} includes both the biases idiosyncratic to study i and the biases θ_j common across empirical studies when applied to future climate change. λ_{ij} is drawn from a normal distribution centered on its group's shared biases θ_j and with standard deviation τ_j. Let $\hat{z}_i$ be the best estimate for empirical study i with $\tilde{z}_i$ as the standard error of that estimate, where the estimates and standard errors are as reported in the main text.

Finally, for coupled models' predictions, define σ_j to be the standard deviation of a study's idiosyncratic bias conditional on its shared biases. Each coupled model i generates "observations" of its central feedback estimate M_{ij} by combining its output with a radiative kernel h as described in Soden et al. (2008). We denote these observations by y_{hi} and let ϕ_j be the standard deviation of those observations around M_{ij}. The standard deviation ϕ_j therefore controls intra-study variation while σ_j controls variation between models. Similarly, τ_j controls variation between empirical studies while $\tilde{z}_i$ describes variation within a single empirical study's estimate.

Table 2.5: The prior distributions used for model parameters and plotted in Figure 2.6. HC(x) is a half-Cauchy distribution with scale parameter x, and t(x, y, z) is a t distribution with location parameter x, scale parameter y, and shape parameter z. See Lemoine (2010a) for more information.

f_{cc} [a]	θ_j [b]	τ_j	σ_j	ϕ_j
t(0,0.15,2)	t(0,0.05,2)	HC(0.01)	HC(0.1)	HC(0.1)

[a] Censored so that values are observed to be less than 1.
[b] Censored so that values are observed to be between -0.5 and 0.5.

The model can be written as:

$$\lambda_{ij} \sim N(\theta_j, \tau_j) \tag{2.10}$$
$$\hat{z}_i \sim t(f_{cc} + \lambda_{ij}, \tilde{z}_i, df) \tag{2.11}$$
$$M_{ij} \sim N(f_{cc} + \theta_j, \sigma_j) \tag{2.12}$$
$$y_{hi} \sim N(M_{ij}, \phi_j) \tag{2.13}$$

where $N(\mu, \sigma)$ is a normal distribution with mean μ and standard deviation σ and where t(x, y, z) is a t distribution with location parameter x, scale parameter y, and shape parameter z. df is the models' degrees of freedom and is equal to 520 for the base case specifications. The prior distributions are given in Table 2.5 and plotted in Figure 2.6, and they follow those used in Lemoine (2010a). The posterior distributions were sampled using Markov chain Monte Carlo methods as implemented in WinBUGS version 1.4.3 (Lunn et al., 2000). Each posterior distribution generated one million samples after a burn-in period of one million samples. The sample size was large enough for multiple chains to converge on the posterior distributions.

Figure 2.7 shows the influence of models' predictions and empirical estimates on the joint distribution for the true feedback factor f_{cc} and the coupled models' shared bias term θ_1 (where the coupled models are group 1). Data from the coupled models can only constrain the sum $f_{cc} + \theta_1$, leading to a ridge in the joint posterior distribution running along values of f_{cc} and θ_1 that produce the same value for $f_{cc} + \theta_1$ and have similar prior densities (Figure 2.7b). However, including the base case empirical results from this paper can further constrain the distribution for f_{cc} because θ_2 is assumed to be independent of θ_1 and the empirical estimates are similar to the coupled models' predictions. A posterior distribution produced using both types of data still has a ridge along similar values of $f_{cc} + \theta_1$, but the ridge is now shorter because the posterior distribution of f_{cc} is also constrained by the

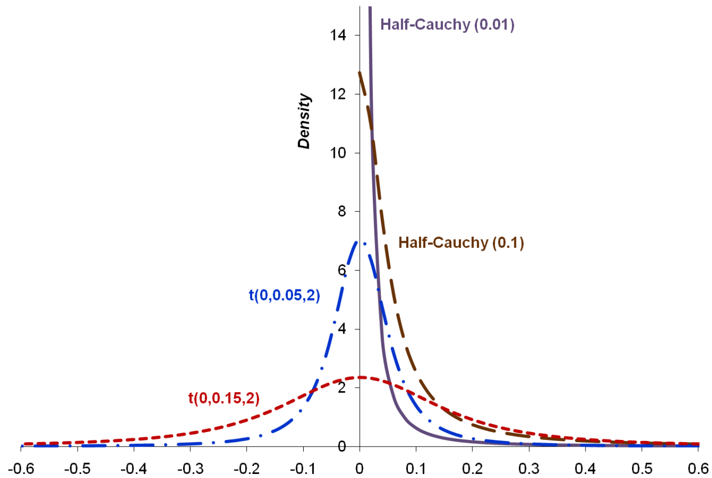
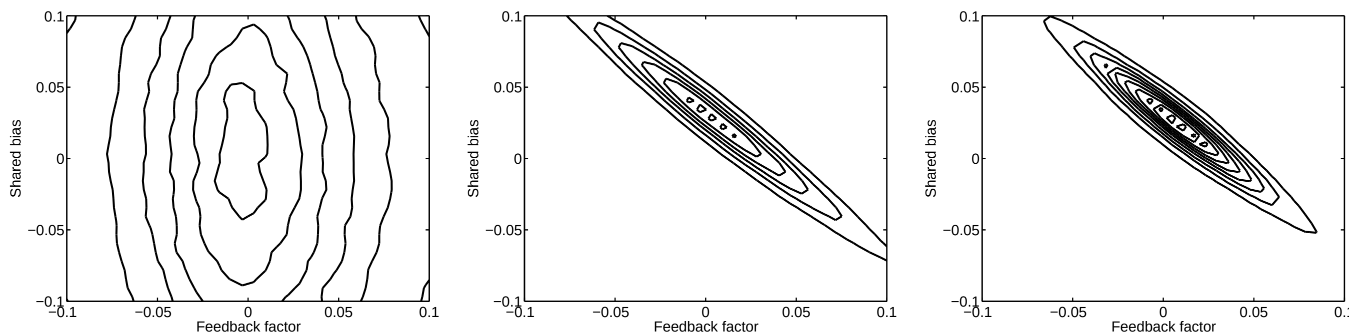


Figure 2.6: The four types of prior distributions described in Table 2.5.

empirical studies' information about the sum $f_{cc} + \theta_2$ (Figure 2.7c).



(a) Prior distributions

(b) Updated from coupled models

(c) Updated from coupled models and empirical studies

Figure 2.7: Contour plots for the joint distribution of the feedback factor f_{cc} (x-axes) and the coupled models' shared bias term θ_1 (y-axes).

Chapter 3

Tipping points and ambiguity in the integrated assessment of climate change[1]

The threat of crossing tipping points in the climate system often serves as an argument for more stringent greenhouse gas emission reductions. We introduce such regime shifts into a recursive relative of the DICE integrated assessment model for establishing optimal climate policies. Each period's carbon dioxide concentration determines the probability of crossing a tipping point, and the policymaker re-optimizes once a tipping point occurs. The probability, timing, and knowledge of tipping points are endogenous. Our policymaker can also display ambiguity aversion in assessing tipping point uncertainty. We find that tipping points can increase the near-term social cost of carbon by more than 50% when they raise climate sensitivity or make damages more convex. They have less of an effect when they increase the atmospheric lifetime of CO_2 or the quantity of non-CO_2 greenhouse gases. Allowing the policymaker to be differentially averse to consumption fluctuations over time and over risk increases the near-term social cost of carbon by 150%, with tipping point possibilities then increasing it by another 50%. The possibility of tipping points is more important for the social cost of carbon than is the ambiguity attitude used in their evaluation. Finally, if the temperature threshold is known to be at 2.5°C or higher, then the policymaker will prevent the tipping point from occurring.

[1]This work owes much to my co-author Christian Traeger.

3.1 Introduction

The threat of climate tipping points motivates calls for aggressive near-term emission reductions to prevent global average temperature from increasing by 2°C relative to pre-industrial levels (e.g., Hansen et al., 2008; Ramanathan and Feng, 2008; Rockström et al., 2009). Tipping points—or irreversible changes in the climate system caused by increasing carbon dioxide (CO_2) and temperature—are poorly understood, difficult to model, and of increasing concern (Alley et al., 2003; Overpeck and Cole, 2006; Smith et al., 2009). By connecting today's climate policy decisions to expectations of future wealth and warming, integrated assessment models (IAMs) estimate the social cost of carbon for use in evaluating emission policies and in cost-benefit analyses (Greenstone et al., 2011). However, because they have not included tipping point possibilities in their system dynamics, these IAMs might consistently underestimate the social cost of carbon.

We answer three questions about the implications of tipping points for IAMs' policy recommendations. First, how do different kinds of tipping points affect the social cost of carbon and optimal abatement policy? Second, if the temperature trigger is known, how high must it be in order to be worth avoiding? Third, is ambiguity aversion with respect to tipping point uncertainty important for the social cost of carbon? We find that tipping points which increase the response of temperature to CO_2 or the damages incurred by high temperatures are more important than tipping points which increase non-CO_2 greenhouse gases or the atmospheric lifetime of CO_2. Optimal policy avoids the temperature trigger if it is known to be at 2.5°C or above, but the cost of avoiding a lower trigger can outweigh the benefits. Finally, aversion to the ambiguity arising from lack of knowledge about tipping point possibilities only mildly affects near-term optimal policy.

To answer these questions, we extend a standard IAM to include endogenous regime shifts, learning about the threshold that triggers a regime shift, and a welfare evaluation based on a generalization of the smooth ambiguity model. Our first extension incorporates tipping points into a recursive version of the Dynamic Integrated model of Climate and the Economy (DICE), a welfare-optimizing IAM (Nordhaus, 2008). A tipping point occurs upon crossing some temperature threshold. Each tipping point irreversibly changes the climate system from its conventional representation in DICE to a new regime that depends on the type of tipping point under consideration. Emission decisions affect whether tipping points occur by determining the temperature expected in each period. Once a tipping point occurs, the global decision-maker optimizes emission and consumption decisions for the post-threshold world. The timing, probability, and welfare consequences of a regime switch are therefore endogenous because they depend on the policies chosen before and after the threshold occurs.

We explore four tipping points (Table 3.1), whose effects are illustrated by the thick arrows in Figure 3.1.[2] The first tipping point changes the response of temperature to increased

[2]We do not model tipping points that might increase welfare because beneficial tipping points are not

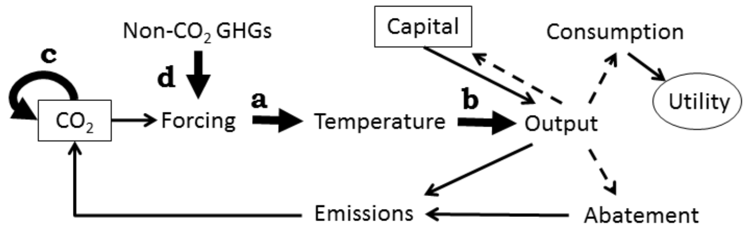


Figure 3.1: A simplified schematic of the modeled relation between the economy and the climate. Dashed arrows indicate the decision variables of consumption, investment, and abatement, and boxes indicate stock variables. Tipping points alter the relationships shown with thick arrows and are labeled as: a) increased climate sensitivity, b) increased convexity of damages, c) weakened CO_2 sinks, and d) increased non-CO_2 forcing.

CO_2 concentrations. More specifically, it increases climate sensitivity, which is the equilibrium warming produced by doubling CO_2 concentrations. This change could occur if land ice sheets begin to retreat on a shorter-than-expected timescale. The second tipping point affects the relation between temperature and output. It increases the convexity of the damage function, which determines the percentage of world GDP that is lost when temperature increases. The convexity of the damage function can be increased by abrupt events that raise sea level (as in the collapse of the West Antarctic or Greenland ice sheets) or that produce unexpectedly nonlinear responses in agricultural production (as in sudden shifts in rainfall patterns). The third tipping point reflects the possibility that carbon sinks weaken beyond the predictions of the simple carbon cycle model represented in DICE. These weakened sinks decrease the decay rate of CO_2, which in turn increases the time for which emitted CO_2 affects the atmosphere. Finally, the fourth tipping point produces a permanent increase in forcing from non-CO_2 greenhouse gases. This corresponds to a large, sustained release of methane from melting permafrost or subsea hydrates, which would raise greenhouse gas concentrations independently of further economic activity.

Most IAMs assume that the dynamics governing the climate and the economy evolve smoothly over time.[3] However, the tipping points described above can be relatively abrupt (Hansen et al., 2008; Lenton et al., 2008). The omission of tipping points has long been recognized as potentially important (e.g., Nordhaus, 1993; Hall and Behl, 2006). One form of sudden change is a generic catastrophic impact that reduces every future period's utility (Gjerde et al., 1999). In contrast, we model sudden temperature-induced changes as altering

much discussed in the climate science literature. We also do not model tipping points that bring the planet to a cold equilibrium (Zaliapin and Ghil, 2010). An interesting analogue of a beneficial tipping point is a new technology that changes the dynamics governing abatement cost.

[3]In DICE-2007, some abrupt damages are included in the calculation of willingness-to-pay used to parameterize the damage function (Nordhaus, 2008), but the dynamics of passing into catastrophic scenarios are not modeled.

Table 3.1: The four tipping points that can occur from crossing a temperature threshold. Each model run uses only one of these post-threshold regimes.

Threshold event	Modeled post-threshold regime
Land ice sheets retreat faster than expected	Climate sensitivity increased from 3°C to 6°C
West Antarctic ice sheet collapses	Damages become a cubic function of temperature
CO_2 sinks weaken	CO_2 decay reduced by 75%
Warming releases methane from hydrates	Non-CO_2 forcing increased by 1.5 W m^{-2}

underlying system dynamics. Further, unlike the more restricted policy setting of Lempert et al. (1994), we embed these shifts in a model of dynamic policy that optimizes welfare at each timestep. Both before and after crossing the threshold, our decision-maker controls consumption, CO_2 abatement, and, via residual output, investment.

Our second extension recognizes that the temperature threshold for a tipping point is unknown. The decision-maker learns about the temperature threshold by observing whether a threshold has or has not been crossed as the world reaches higher temperatures. The chosen emissions determine the probability that a tipping point occurs, and this probability itself depends on the temperature produced by previous emission decisions. Increasing global temperature produces a greater chance of a tipping point occurring when temperature is already high. Both the probability of a tipping point occurring and the decision-maker's knowledge of the probability distribution are endogenous in the sense that they depend on chosen emission levels. Similarly to our work, Keller et al. (2004) extended DICE to endogenously determine when a threshold is potentially crossed. They considered an abrupt collapse of the thermohaline circulation (or Gulf Stream) that permanently shifts the damage function. Our model differs by including pervasive temperature stochasticity and by having the decision-maker endogenously learn about the threshold's location through the chosen emission path.

Several studies have analyzed uncertainty in DICE by drawing model parameters from a distribution and then determining optimal policy under certainty for each realization (e.g., Nordhaus, 2008; Ackerman et al., 2010). They approximated the optimal policy under uncertainty as the average policy from the various deterministic model runs. The resulting policy is usually not the same as optimal policy under uncertainty (Newbold and Daigneault, 2009). Other models have optimized consumption in the face of persistent stochasticity but imposed exogenous greenhouse gas policies (Gerst et al., 2010). We instead convert DICE into a recursive dynamic programming model (e.g., Kelly and Kolstad, 1999; Leach, 2007; Crost and Traeger, 2010). This enables us to analyze optimal policies under uncertainty about temperature change and about the temperature threshold that triggers a regime shift.

Each period's optimal policies reflect the possibility that the next period's temperature will not be as expected and, moreover, that a climate threshold will be crossed.

By extending a full numerical IAM, we estimate the effect of uncertain tipping points on the social cost of carbon. Other relevant work has investigated the effects of uncertain climate tipping points in less complex models with different policy problems. It has considered how the possibility of a climate catastrophe affects the optimal steady state CO_2 concentration (Tsur and Zemel, 1996), how the endogenous risk of especially high-damage outcomes affects irreversible investment in abatement capital (Fisher and Narain, 2003), and how hyperbolic discounting increases the salience of a low probability, distant catastrophe (Karp and Tsur, 2007). Another pair of studies considered optimal control problems in the presence of uncertain thresholds that cause a stream of disutility (Nævdal, 2006; Nævdal and Oppenheimer, 2007). Finally, if emissions cannot be fine-tuned but instead are controlled only by discrete policies of predefined magnitude, then the policymaker must choose when to adopt the emission policies. This timing depends in part on how their adoption affects the possibility of catastrophe (Baranzini et al., 2003; Guillerminet and Tol, 2008).

Beyond the climate context, other work has considered how an endogenous, uncertain threshold that makes pollution stop decaying altogether can produce multiple optimal paths for consumption and environmental quality (Kama et al., 2011). This threshold's effect on system dynamics is similar to our tipping point with weakened CO_2 sinks. A separate effort analytically modeled a shift in pollutant loading upon crossing a stochastic, reversible threshold (Brozović and Schlenker, 2011), whereas we consider a fixed, irreversible threshold and model learning about its location. Their main result was that precaution is non-monotonic in the variance of the distribution for the threshold level: precaution at first increases with the variance as pollution levels just below the expected threshold have a greater and greater chance of crossing it, but precaution eventually decreases when the variance becomes especially high because more probability mass then falls on threshold values that are either too high or too low to warrant precaution.[4] The possibility of collapse or altered system dynamics also arises in discussion of the optimal use of renewable resources. Polasky et al. (2010) considered a regime shift that reduces the resource's growth function. They found that the pre-threshold policy becomes more precautionary when the possibility of the regime shift is endogenous. Our work also examines the implications of an endogenous regime shift that affects system dynamics, but we do so with a nonlinear utility function that allows for risk aversion and we explore several different types of regime shifts.

Our third extension of the standard IAM distinguishes different types of uncertainty when evaluating welfare. Specifically, our decision-maker can be more averse to poorly understood tipping point uncertainty than to better understood temperature risk. We use "objective risk" to describe uncertain outcomes whose probabilities can be gleaned from data. In our

[4]Because we have exogenous variables evolving in time, one difference between our model and that of Brozović and Schlenker (2011) is that the cost of the policies needed to stay just below a known threshold also changes over time.

model, this describes temperature stochasticity. We use the term "subjective uncertainty" to describe uncertain outcomes when there is less information available for determining probabilities. This deficiency in probabilistic knowledge applies to tipping points, which are more poorly understood than other climate phenomena (Alley et al., 2003; Lenton et al., 2008; Ramanathan and Feng, 2008; Kriegler et al., 2009; Smith et al., 2009). This distinction in types of uncertainty reaches back to Keynes (1921: Chapter VI) and Ellsberg (1961), who add an additional confidence weight to probabilities in order to describe uncertainty more comprehensively. We employ a recent extension by Traeger (2010b) of the Klibanoff et al. (2005, 2009) smooth ambiguity model in order to capture the decision-maker's ambiguity attitude, or the decision-maker's attitude to varying confidence in probability judgments. The model is consistent with conventional decision-theoretic axioms, including time consistency and a minimally modified version of the von Neumann and Morgenstern (1944) axioms. Lange and Treich (2008) have applied the Klibanoff et al. (2005) model to analytically explore the implications of ambiguity about a climate damage parameter. Millner et al. (2010) applied the multiperiod evaluation function of Klibanoff et al. (2009) to evaluate exogenous consumption paths from DICE simulations under ambiguity about climate sensitivity. They explored the response of this evaluation to changes in ambiguity attitude. Finally, Hennlock (2009) applied robust control theory to a two-sector analytic version of DICE. The robust control model can be interpreted as the limiting case of extreme ambiguity aversion in our smooth ambiguity framework. Our work differs from these in numerically implementing a model of ambiguity aversion that is not only used to generate endogenous policy paths but also entails optimal updating of the ambiguous distribution.

Section 3.2 explains how we extend DICE, emphasizing the inclusion of tipping points and the welfare evaluation that disentangles three types of aversion. In section 3.3, we discuss the details of our four tipping point scenarios. Section 3.4 presents the results for each tipping point with a known threshold, with an unknown threshold, and with ambiguity aversion. We conclude in section 3.5 with a discussion of the implications for climate science, economics, and policy. The two appendices describe the model calibration and equations.

3.2 Introducing tipping points and ambiguity aversion into DICE

Our infinite horizon model extends Crost and Traeger (2010), which is a stochastic dynamic programming relative of the DICE-2007 model by Nordhaus (2008). DICE is a Ramsey-Cass-Koopmans growth model that has an aggregate world economy interacting with a climate module. Gross economic output (or potential GDP) is determined by an endogenous capital stock, an exogenously growing labor force, and exogenously improving production technology. Gross output produces CO_2 emissions. Non-abated CO_2 emissions accumulate in the atmosphere and ultimately translate into global warming, which causes damage proportional to world output. The control variables of the model are abatement and consumption, and residual output not allocated to these two options becomes capital investment (Figure 3.1). We implement a recursive modeling structure for the following reasons. First, we can account for temperature stochasticity. Second, we can account for tipping point uncertainty and endogenous regime shifts. Third, we can employ a more comprehensive approach to welfare evaluation that distinguishes between aversion to consumption varying over time, aversion to temperature risk, and aversion to ambiguous tipping point occurrences. In order to replicate DICE in a stochastic dynamic programming framework, and in order to avoid falling victim to the curse of dimensionality, we limit our model to the state variables of capital, the stock of CO_2 in the atmosphere, and time. For this purpose, we approximate the carbon cycle and the cooling effect of ocean heat capacity as interpolated functions of the CO_2 stock and time rather than as additional state variables. Moreover, we translate DICE's intrinsic warming delay into increased strength of the cooling reservoirs.[5] The two appendices provide more details on the equations governing the economy and the climate system as well as on the calibration and interpolation procedures.

In the following we explain the richer uncertainty and welfare structure of our model. We start by reviewing the welfare function in the absence of tipping points or ambiguity. For a given consumption path, or, more precisely, for a given probability tree over future consumption, we represent aggregate welfare in period t by the function V_t. It is defined recursively by aggregating in every period t immediate utility $u(c_t) = \frac{c_t^{1-\eta}}{1-\eta}$ from per capita consumption c_t with uncertain future welfare $\tilde{V}_{t+1}$:[6]

$$
\begin{aligned}
V_t(c_t, \tilde{V}_{t+1}) =& u(c_t) + \beta_t \; f_o^{-1} \left[\; \mathrm{E}_t \; f_o \left[\tilde{V}_{t+1}(c_{t+1}, \tilde{V}_{t+2}) \right] \right] && (3.1) \\
=& \frac{c_t^{1-\eta}}{1-\eta} + \frac{\beta_t}{1-\eta} \left\{ \int \left[(1-\eta) \tilde{V}_{t+1}(c_{t+1}, \tilde{V}_{t+2}) \right]^{\frac{1-\gamma_o}{1-\eta}} d\mathbb{P}_t \right\}^{\frac{1-\eta}{1-\gamma_o}} . && (3.2)
\end{aligned}
$$

Future welfare is discounted with the factor β_t. The discount factor captures a rate of pure time preference of 1.5% as in DICE-2007. It also adjusts for population growth so that equation (3.2) effectively measures welfare as the population-weighted sum of instantaneous per capita utility.[7] When a consumption path is certain, the above equation collapses to the standard discounted sum of utility over time with constant intertemporal elasticity of substitution of η^{-1}. Thus, η captures the preference for consumption smoothing over time.

[5]Including DICE's delay equation for temperature would increase the number of state variables needed. We instead adjust the transient feedback (or cooling reservoir) to calibrate our model without tipping points or uncertainty to DICE's optimized output.

[6]We use the tilde to denote random variables.

[7]We originally have an instantaneous utility function $L_t u\left(\frac{C_t}{L_t}\right)$ as in DICE. For convenience of representation and of numerical solution, we divide through with the population L_t and represent welfare as a function of per capita consumption $c_t = \frac{C_t}{L_t}$. The discount factor then has to pick up the exogenous change in population on top of pure time preference.

Greater η means reduced intertemporal elasticity of substitution and greater aversion to consumption varying over time. In equation (3.2), welfare in period $t+1$ is uncertain only because of climatic fluctuations affecting the temperature realized in a given period. This temperature is determined by the realization of a shock coming from an independent and identical distribution having probability measure $\mathbb{P}$. The stochastic shock multiplies climate sensitivity in the equation determining temperature change (see Section 3.3 and the first appendix), producing a lognormally distributed random variable with a mean of 2.8 and a variance of 0.5. This variable's distribution is similar to those given in Meehl et al. (2007) for the climate sensitivity parameter.[8] The multiplicative noise captures period-to-period temperature variability that makes extreme outcomes more likely as CO_2 increases. The inclusion of uncertainty means that we have to take the expectation over welfare in the second period. A concave function f_o determines aversion with respect to welfare uncertainty. Traeger (2010a) gives an axiomatic characterization of the function f_o, labeling it intertemporal risk aversion. Under intertemporal risk neutrality, or neutrality with respect to welfare uncertainty, the function f_o is linear and the expression (3.2) reduces to the intertemporally additive standard model that can be defined non-recursively. We adopt a power function for characterizing intertemporal aversion to objective risk: $f_o(V) = ((1-\eta)V)^{\frac{1-\gamma_o}{1-\eta}}$. This type of preference relation is known as generalized isoelastic Weil (1990) or Epstein and Zin (1989) preferences. The function $f_o \circ u(c_t) = \frac{c_t^{1-\gamma_o}}{1-\gamma_o}$ characterizes Arrow-Pratt relative risk aversion γ_o.

Uncertainty in equation (3.2) only considers the period-to-period variability in temperature, not the chance of crossing a climate threshold. We now describe the value functions solved for in the model. Crossing the threshold causes an irreversible switch from a pre-threshold regime (indicated by $\psi_t = 0$) to a post-threshold regime (indicated by $\psi_t = 1$). A regime switch immediately after period t means that $\psi_s = 0 \ \forall s \leq t$ and $\psi_s = 1 \ \forall s > t$. Once the regime switch occurs, all that matters for the decision problem are the post-threshold equations of motion and the state variables, which are capital k_t (per effective unit of labor), the stock M_t of CO_2 in the atmosphere, and time t. The policymaker optimizes over consumption c_t and the abatement rate μ_t so as to maximize the welfare function (3.2) under the constraints of the equations of motion governing the economy and the climate system. This gives the dynamic programming equation:[9]

$$V_{\psi=1}(k_t, M_t, t) = \max_{c_t,\mu_t} \frac{c_t^{1-\eta}}{1-\eta} + \frac{\beta_t}{1-\eta} \left\{ \int [(1-\eta)V_{\psi=1}(k_{t+1}, M_{t+1}, t+1)]^{\frac{1-\gamma_o}{1-\eta}} d\mathbb{P} \right\}^{\frac{1-\eta}{1-\gamma_o}}, \tag{3.3}$$

where the state variables for time $t+1$ derive from the equations of motion for the economy

[8]We implement the continuous distribution numerically using a Gauss-Legendre quadrature rule with 8 nodes.

[9]As in Crost and Traeger (2010), the model is actually solved using a transformation mapping the infinite time horizon to the unit interval. Compare also Kelly and Kolstad (1999).

and the climate system described in the second appendix. As discussed below, we solve the dynamic programming problem for the value function $V_{\psi=1}$ numerically for four different post-threshold regimes as well as for different preference parameter specifications. We use a function iteration approach employing Chebychev polynomials for the value function approximation at a set of Chebychev nodes in the three-dimensional state space.

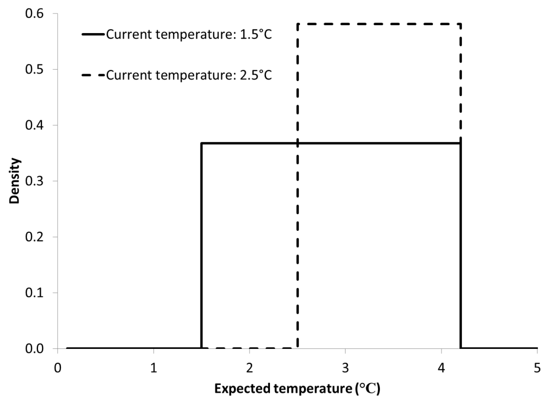
Once we have solved for the post-threshold value function, we use it to solve the pre-threshold value function. We assume that the system passes from the pre-threshold regime ($\psi_t = 0$) into the post-threshold regime ($\psi_{t+1} = 1$) if the expected temperature change $E_t[T_{t+1}]$ relative to pre-industrial levels crosses a threshold. We make the threshold depend on expected temperature rather than on actual temperature realizations because, first, expected temperature captures the intuition that medium-term changes in temperature are more likely to trigger tipping points and, second, it saves a state variable.[10] When the threshold temperature is known to be T^*, we call it a "certain threshold."

In general, however, the decision-maker does not know where the threshold lies, making the threshold temperature a random variable $\tilde{T}$. The probability distribution for $\tilde{T}$ follows from assuming that a tipping point is sure to occur by the time the world reaches a temperature $\bar{T}$ and that any temperature between present levels and $\bar{T}$ has an equal chance of being the threshold. Further, we assume that the time 0 expected value for the threshold is 2.5°C: $\mathrm{E}_0\, \tilde{T} = 2.5$°C, which combines with the assumed uniform distribution for $\tilde{T}$ to give $\bar{T} = 4.22$°C (Figure 3.2a). This value for $\mathrm{E}_0\, \tilde{T}$ is consistent with the political 2°C limits for avoiding dangerous anthropogenic interference. Further, in Smith et al. (2009), 2.5°C is in the upper end of the temperature change that produces significant risk of large-scale discontinuities and is just below the temperature that produces severe risk. In the baseline optimal policy scenario without a threshold, temperature rises above $\mathrm{E}_0\, \tilde{T}$ in the year 2081 and never reaches $\bar{T}$. The probability of crossing the threshold between periods t and $t+1$ conditional on not having crossed the threshold by time t is:

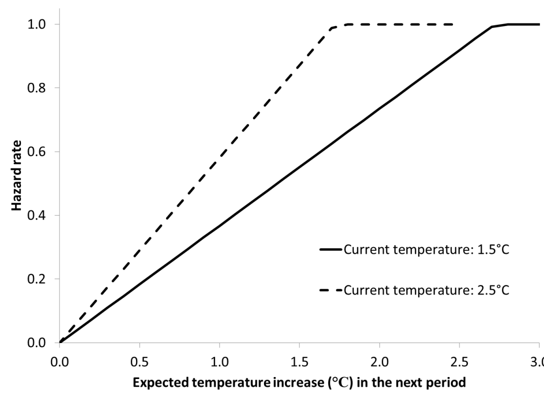
$$h(E_t[T_{t+1}]) = \max\left\{0, \frac{\min\{E_t[T_{t+1}], \bar{T}\} - E_{t-1}[T_t]}{\bar{T} - E_{t-1}[T_t]}\right\} . \tag{3.4}$$

This expression is the hazard rate. The time t expectation of temperature at time $t+1$ is conditional on the CO_2 stock and the transient feedback at time $t+1$ (see the first ap-

[10] In DICE-2007, the CO_2 stock increases monotonically until the model reaches a sufficiently high level of abatement. From this point on, the decay rate of CO_2 outweighs the flow of emissions, making the CO_2 stock decrease monotonically. With an increasing CO_2 stock, the probability of crossing the threshold is proportional to the difference between expected temperature in the next period and the expected temperature for the current period (as determined by the current CO_2 stock). For a decreasing CO_2 stock, the probability of crossing the threshold is 0. As long as expected temperature is a quasiconcave function of time, we do not need an additional state variable to keep track of the highest historic expected temperature. As in DICE-2007, CO_2 concentrations in our model follow a path that is quasiconcave in time, and the resulting expected temperature is also quasiconcave in time because it is determined by the CO_2 stock and a transient feedback that itself generally increases with time (see the first appendix).



(a) Probability distribution for $\tilde{T}$



(b) Hazard rate

Figure 3.2: As the time t expected temperature increases without crossing a threshold, the probability distribution for the threshold level $\tilde{T}$ places more mass on temperatures yet to be reached. Each additional increase in temperature therefore also produces a greater risk of crossing the threshold (i.e., the hazard rate is greater for higher $E[T_t]$).

pendix).[11] These are known at time t, and expectations are formed over the noise term generating stochastic temperatures. The hazard rate for further temperature increases is greater when the current temperature is greater (Figure 3.2b). The hazard rate is endogenously determined by the decisions that are optimal for each state of the world.

Expected welfare in the pre-threshold regime is the sum of the known current welfare and the discounted uncertain future welfare. If the threshold is not crossed during the next period, welfare is given again by the pre-threshold welfare function $V_{\psi=0}$ (evaluated at the next period's states and subject to the temperature uncertainty captured by $\mathbb{P}$). However, with a probability given by the hazard rate, next period's welfare is determined by the post-threshold regime in the function $V_{\psi=1}$. Therefore, expected welfare in the next period is the weighted average of welfare with and without crossing the threshold, where the weights

[11] Our uncertain threshold is therefore distributed in the two-dimensional state space of time and the CO_2 stock (compare Nævdal, 2006; Nævdal and Oppenheimer, 2007).

follow from the hazard rate. This gives the pre-threshold dynamic programming equation:

$$\begin{aligned} V_{\psi=0}(k_t, M_t, t) &= \max_{c_t,\mu_t} u(c_t) + \beta_t f_o^{-1}\left[\int (f_o \circ f_s^{-1})\Big\{[1-h(E_t[T_{t+1}])]f_s[V_{\psi=0}(k_{t+1}, M_{t+1}, t+1)] \right. \\ &\qquad \left. + h(E_t[T_{t+1}])f_s[V_{\psi=1}(k_{t+1}, M_{t+1}, t+1)]\Big\} d\mathbb{P}\right] \\ &= \max_{c_t,\mu_t} \frac{c_t^{1-\eta}}{1-\eta} + \frac{\beta_t}{1-\eta}\left[\int \Big\{[1-h(E_t[T_{t+1}])]\big[(1-\eta)V_{\psi=0}(k_{t+1}, M_{t+1}, t+1)\big]^{\frac{1-\gamma_s}{1-\eta}} \right. \\ &\qquad \left. + h(E_t[T_{t+1}])\big[(1-\eta)V_{\psi=1}(k_{t+1}, M_{t+1}, t+1)\big]^{\frac{1-\gamma_s}{1-\eta}}\Big\}^{\frac{1-\gamma_o}{1-\gamma_s}} d\mathbb{P}\right]^{\frac{1-\eta}{1-\gamma_o}} . \end{aligned} \tag{3.5}$$

The function f_s characterizes intertemporal aversion to subjective risk, which in our case is the chance that a tipping point is crossed. As with f_o, we adopt a power function $f_s(z) = f_s \equiv [(1-\eta)z]^{\frac{1-\gamma_s}{1-\eta}}$, $z \in \mathbb{R}_+$. In a one-commodity setting, γ_s can be understood as a measure of Arrow-Pratt relative risk aversion with respect to subjective uncertainty (or to poorly understood uncertainty). As described earlier, the probability of crossing the threshold before the next period is generally less confidently known than is the distribution for temperature in a given period. For this reason, we have $\gamma_s \geq \gamma_o$, allowing the decision-maker to be more averse to the chance of a threshold crossing (determined by $h(\cdot)$) than to the risk produced by not knowing the next period's temperature exactly (determined by $\mathbb{P}$). The function $f_{amb}(z) = (f_s \circ f_o^{-1})(z) = z^{\frac{1-\gamma_s}{1-\gamma_o}}$, $z \in \mathbb{R}_+$ captures an extended form of what the literature calls smooth ambiguity aversion (Traeger, 2010b).[12] We solve for the pre-threshold value function employing the same method as described above for the post-threshold value function.

The baseline parameterization of aversion to temperature risk and to intertemporal consumption fluctuations (labeled "partially disentangled preferences") matches the utility function in DICE-2007, with Arrow-Pratt relative risk aversion γ_o and aversion to intertemporal substitution η both equal to 2 (Table 3.2). Runs with this parameterization do not disentangle preferences over time from preferences over the temperature change lottery. To assess the effect of ambiguity aversion, we consider a case with ambiguity neutrality ($\gamma_s = 2$, which implies that f_{amb} is linear), a moderate level of ambiguity aversion close to the calibration result by Ju and Miao (2009) in the asset pricing context ($\gamma_s = 9.5$), and a bounding case of extreme ambiguity aversion ($\gamma_s = 50$). In addition, we include simulations with "fully disentangled preferences" to assess their implications for thresholds' effects and ambiguity

[12]The model of smooth ambiguity aversion goes back to Klibanoff et al. (2005, 2009). However, these authors only allow a twofold disentanglement of preferences that distinguishes aversion to ambiguous distributions but does not separate standard risk aversion from aversion to intertemporal consumption fluctuations. Traeger (2010b) gives an axiomatic extension of the concept of smooth ambiguity aversion to the current case of a threefold disentanglement.

Table 3.2: The welfare specifications used to assess the effect of disentangling preferences about consumption over time from preferences about consumption over uncertain states. The standard preference structure from DICE-2007 in our setting with uncertainty about temperature and tipping points has $\eta = \gamma_o = \gamma_s = 2$.

Preference disentanglement	Aversion to intertemporal substitution (η)	Aversion to temperature risk (γ_o)	Aversion to tipping point uncertainty (γ_s)
Partial	2	2	{2, 9.5, 50}
Full	2/3	9.5	{9.5, 19, 50}

aversion. Here, we reduce the aversion to intertemporal substitution to $\eta = 2/3$ and use Arrow-Pratt relative risk aversion $\gamma_o = 9.5$ (Vissing-Jørgensen and Attanasio, 2003). Recent work has used a similar preference parameterization to explain equity risk premia and cyclical risks (Bansal and Yaron, 2004; Bansal et al., 2010). In this fully disentangled model, we consider ambiguity neutrality ($\gamma_s = 9.5$), a moderate level of ambiguity aversion ($\gamma_s = 19$), and an extreme level of ambiguity aversion ($\gamma_s = 50$).

3.3 Modeled tipping points

We now further describe the four tipping points listed in Table 3.1 and illustrated in Figure 3.1. The second appendix shows how optimal policy responds to the modeled changes in system dynamics. It helps to first introduce the equations describing equilibrium temperature change, damages, and the evolution of the CO_2 stock. Equilibrium temperature change T_{equil} relative to pre-industrial levels is given by:

$$T_{equil} = s \left[\frac{\ln(M_t/M_{pre})}{\ln 2} + \frac{EF_t}{5.35 \ln 2} \right] . \tag{3.6}$$

The parameter s is climate sensitivity, M_t and M_{pre} are the time t and pre-industrial CO_2 stocks, and EF_t is the exogenous time t non-CO_2 forcing. To obtain time t temperature change T_t, we multiply s by the lognormally distributed stochastic shock described above and then adjust T_{equil} for the cooling reservoir described in the first appendix. The expected value of T_t increases with T_{equil}. Let Y_{gross} be the total output produced by the time t capital stock. Then the output available for allocation to consumption, abatement, and investment is $Y_{gross}/(1 + D)$. The function D gives damages due to temperature change:

$$D = b_2 T_t^{b_3} . \tag{3.7}$$

The parameter b_2 equals 0.0028388 in DICE-2007. The damage function in DICE is quadratic, giving $b_3 = 2$. Finally, the CO_2 stock M_t evolves according to the following transition equation:

$$M_{t+1} = M_{pre} + (1 - \phi_t')(M_t - M_{pre}) + E_t , \tag{3.8}$$

where E_t gives time t emissions. The term ϕ'_t is defined in the first appendix. It determines the decay of the CO_2 stock towards pre-industrial levels. It is a function of time and the CO_2 stock and is inferred from the output of DICE-2007.

The first tipping point occurs if we observe land ice sheets to be retreating over shorter-than-expected timescales in response to an experienced increase in temperature. Common definitions of climate sensitivity assume that these "slow feedbacks" do not affect temperature on relevant timescales. If they do begin to operate, they amplify the warming predicted using conventional estimates of climate sensitivity. This tipping point shifts our model into a post-threshold regime with increased climate sensitivity $\hat{s}$:

$$\hat{s} = 2s \ . \tag{3.9}$$

Climate sensitivity in the pre-threshold regime is 3°C as in DICE, and post-threshold climate sensitivity becomes 6°C (Hansen et al., 2008).[13] Each unit of emissions causes more temperature change than expected, which quadruples damages from equilibrium temperature change.

The second tipping point occurs when a sudden, irreversible change such as the collapse of the West Antarctic or Greenland ice sheets increases impact assessments for higher temperatures (Oppenheimer, 1998; Vaughan, 2008; Notz, 2009). DICE uses a damage function that is quadratic in temperature, but many have raised concerns about the fit of a quadratic function at high levels of temperature change (e.g., Wright and Erickson, 2003; Ackerman et al., 2009; Newbold and Daigneault, 2009; Weitzman, 2009; Ackerman et al., 2010; Hanemann et al., 2010). This tipping point increases the convexity of the damage function by changing the exponent on T_t to $\hat{b}_3$. We parameterize this as making damages a cubic function of temperature in the post-threshold regime:[14]

$$\hat{b}_3 = b_3 + 1 = 3 \ . \tag{3.10}$$

Damages are now multiplied by an additional factor equal to current temperature. Along our model's optimal paths, the expectation of this factor is always less than the quadrupling that would be caused by the slow feedback tipping point with equilibrium temperature. However, unexpectedly high temperature realizations can cause more than a quadrupling in damages.

The third and fourth tipping points model degradation of carbon sinks and and activation of methane sources. The third tipping point reflects the possibility that carbon sinks weaken beyond the predictions of coupled climate-carbon cycle models (Raupach et al., 2008). The many processes through which the climate and carbon cycle affect each other are difficult to model and to calibrate (Luo, 2007), making it hard to rule out extreme outcomes (e.g.,

[13] We also remove temperature stochasticity in this post-threshold regime so as not to confuse the tipping point's effect through the possibility of extremely high temperature realizations.

[14] This regime with more convex damages is similar to the case considered by Azar and Lindgren (2003), but their regime switch happens only in 2035 and with a low probability that is exogenous (i.e., does not depend on emission decisions).

Sitch et al., 2008). Warming-induced changes in oceans (Le Quéré et al., 2007), soil carbon dynamics (Eglin et al., 2010), and standing biomass (Huntingford et al., 2008) could affect the uptake of CO_2 from the atmosphere. In order to represent an extreme form of weakened sinks, we parameterize this tipping point as causing a 75% reduction in the decay rate of CO_2:

$$\hat{\phi}_t = \phi'_t/4 \ . \tag{3.11}$$

The change from ϕ'_t to $\hat{\phi}_t$ increases the time for which a unit of emitted CO_2 affects atmospheric CO_2, which becomes more important as time passes and CO_2 accumulates in the atmosphere to a greater degree than it would have. However, the present value of the damage caused by this effect is reduced by discounting and by the concave relationship between CO_2 and temperature.

The fourth tipping point corresponds to a permanent increase in forcing from non-CO_2 greenhouse gases, as if from a large, sustained release of methane from melting permafrost or subsea hydrates (Hall and Behl, 2006; Archer, 2007; Schaefer et al., 2011). A large methane release is one of the hypothesized triggers for ancient periods of rapid warming (Zachos et al., 2008). We represent this tipping point as permanently increasing each period's non-CO_2 forcing by 1.5 W m^{-2}:

$$\widehat{EF}_t = EF_t + 1.5 \ . \tag{3.12}$$

The 1.5 W m^{-2} additional forcing is at the low end of a range of plausible methane emission rates during ancient warming (Schmidt and Shindell, 2003), and it is equivalent to increasing the CO_2 concentration by 30%. This increase in non-CO_2 greenhouse gases multiplies damages by a factor that is smaller than temperature as long as the current temperature is greater than 2.3°C. It always increases damages by less than does the slow feedback tipping point and, for most modeled values of the temperature threshold, it also increases damages by less than does the damage convexity tipping point.

3.4 Results

We compare several sets of model runs to assess how the social cost of carbon and optimal CO_2 concentrations respond to different tipping points, to uncertainty about the temperature threshold for a tipping point, and to additional aversion to tipping point risk (Table 3.3). Further, we also solve each type of model with partially disentangled preferences and with fully disentangled preferences (Table 3.2). The baseline version of the model is the standard DICE model plus temperature stochasticity. Period-to-period temperature uncertainty has

Table 3.3: The model runs used to assess the effects of climate tipping points and ambiguity aversion. Each of these runs is done with partially disentangled preferences and with fully disentangled preferences.

Model	Description
Base case	No thresholds can occur
Certain threshold	Policy precisely controls when and whether a threshold occurs
Uncertain threshold	Policy controls the probability that a threshold occurs
Ambiguity aversion	Additional aversion to threshold uncertainty

a negligible effect on policy in these runs.[15,16] When unexpected temperature realizations occur, it is the residual output (i.e., investment) that responds, and the anticipated possibility of unexpected outcomes is not important enough to affect the consumption and abatement policies chosen before the temperature variability is resolved. A second set of runs has a tipping point occurring at a known threshold. The decision-maker knows that the world will change once reaching the temperature T^*. She can therefore adjust emissions to delay or avoid crossing the threshold and also adjusts optimally after crossing the threshold. A third set of runs makes the decision-maker uncertain about the temperature threshold that triggers the tipping point. Additional emissions increase the chance of crossing into the new regime, and the decision-maker updates beliefs about the threshold based on whether a tipping point has occurred. Finally, a fourth set of runs includes aversion to tipping point ambiguity, which makes the decision-maker more averse to tipping point uncertainty than to temperature change uncertainty. In each model run with tipping points, the decision-maker only faces one type of tipping point and knows which type that is.

The effect of each tipping point on post-threshold policy and welfare determines how pre-threshold policy responds to awareness of the tipping point. Pre-threshold policy affects the chance of crossing the tipping point and also affects the capital and CO_2 stocks at the time the threshold is crossed. In Figure 3.3, the differences in the time paths within a plot show the effect of tipping point considerations and preference disentanglement on the expected social cost of carbon and on expected CO_2 concentrations. Comparing time paths in plots on different rows shows the effects of different types of tipping points. Figure 3.4 compares

[15]In the baseline runs, uncertainty about temperature change does not affect expected policy by much with either standard or fully disentangled preferences. With standard preferences, this remains true even when we use a model version calibrated to yield higher temperatures and even when we use a thicker-tailed distribution for temperature stochasticity. We generate this thicker-tailed distribution by adding a normally distributed random variable to the feedback factors described in the first appendix (Roe and Baker, 2007). This random variable was calibrated to a reference distribution in Lemoine (2010a). The expected climate sensitivity implied by this random variable and the typical constant feedback is 2.9°C. The thicker-tailed distribution is closer to those used in Newbold and Daigneault (2009) and Gerst et al. (2010).

[16]Temperature uncertainty can noticeably increase the social cost of carbon in post-threshold regimes where higher temperature realizations cause greater damage.

the abatement rate in models with an unknown threshold and standard preferences. The effect of considering a tipping point possibility depends on where the temperature threshold might be: if it is known to be at 2.5°C or higher, the policymaker will ensure it is not crossed (Figure 3.5). Optimal policy may still incur substantial (though reduced) risk of crossing the uncertain threshold (Table 3.4), and a possible tipping point increases the year 2015 social cost of carbon by over 50%, while moving to fully disentangled preferences has an even greater effect (Table 3.5). Finally, moderate and extreme ambiguity aversion increase the social cost of carbon, but only by a small amount until later in the 21st century when the probability of crossing a tipping point becomes higher (Figure 3.6).

The climate sensitivity and damage convexity tipping points have the greatest effect on post-threshold damages and, as a result, also have the greatest effect on pre-threshold policy. The climate sensitivity tipping point quadruples equilibrium damages under expected temperature outcomes, and the damage convexity tipping point multiplies pre-threshold damages by a factor equal to the temperature outcome. This temperature outcome is usually less than 4 but can be greater under unexpected temperature realizations, meaning that less likely temperature outcomes with increased damage convexity can increase damages by more than does the climate sensitivity tipping point. When the threshold is known to be at 2°C or higher, pre-threshold policy avoids it with either type of tipping point. When the threshold is uncertain, the two tipping point possibilities increase the social cost of carbon in similar ways and keep the CO_2 stock on a lower trajectory (peaking around 520 ppm if the world is lucky and the threshold is not crossed) than in the baseline case (which peaks near 650 ppm). Both produce similar abatement policies, generally increasing baseline abatement over the 21st century by more than 60%. With fully disentangled preferences, they decrease the peak CO_2 concentration from around 530 ppm to around 470 ppm. Either uncertain threshold increases the 2015 social cost of carbon by 50% relative to its baseline value under a given preference specification. Extreme ambiguity aversion can increase the social cost of carbon late in the 21st century (when tipping points are more likely) by over 40% relative to ambiguity neutrality. Moderate ambiguity aversion does not have a discernible effect on either optimal policy or on the social cost of carbon.

In contrast, when the tipping point decreases the decay rate of CO_2, neither the possibility of crossing an uncertain threshold nor the fact of having crossed it greatly affects optimal policy. CO_2 does accumulate in the atmosphere to a greater degree after crossing the threshold, but this accumulation takes time, which means discounting reduces the loss from the threshold crossing. Further, while an additional unit of CO_2 will cause more damage by affecting the atmosphere for longer once the threshold is crossed, the additional CO_2 also matters less when concentrations are already high because warming is a concave function of CO_2. Recognizing the possibility of an unknown threshold does increase optimal abatement and the social cost of carbon relative to the baseline case, but this effect is small. The probability of eventually crossing the threshold only decreases by 0.07 when the decision-maker is made aware of threshold possibilities, and neither moderate nor extreme ambiguity aversion has much additional effect.

Finally, when the tipping point increases non-CO_2 forcing and the decision-maker does not know where the threshold is, optimal policy keeps the pre-threshold temperature lower than in the baseline run. This lower peak temperature reduces the total chance of crossing a threshold: if the policymaker is unaware of the threshold possibility (therefore setting policy as in the baseline case), the optimal policy with standard preferences gives about a 65% chance of crossing the threshold, but optimal policy when aware of the threshold possibility reduces the chance of eventually crossing it to about 50%. This is about 5 percentage points higher than the chance of eventually crossing the climate sensitivity and damage convexity tipping points, but it is about 8 percentage points lower than the chance of eventually crossing the decay rate tipping point. With fully disentangled preferences, the effect of this non-CO_2 forcing tipping point is closer to those of the climate sensitivity and damage convexity tipping points: it reduces the 49% chance of crossing the threshold along the baseline path to a 37% chance of eventually crossing it. The preference specification matters greatly here because the increased non-CO_2 forcing raises damages far out into the future even after CO_2 concentrations have returned near to pre-industrial levels, whereas the damage convexity and climate sensitivity tipping points' effects are less prominent once CO_2 has returned to those lower levels. The implicit reduction in discounting in the fully disentangled preference specification therefore makes the non-CO_2 tipping point look more like the two worse ones.[17] Finally, this tipping point has almost as strong an effect on the year 2015 social cost of carbon as do the climate sensitivity and damage convexity tipping points.

To better understand the effect of threshold uncertainty, it helps to consider the cost imposed by an extra unit of CO_2 emissions. This marginal damage (or social cost of carbon) determines the optimal level of abatement, which is set so that marginal cost equals marginal damage. Additional emissions always increase the next period's stock of CO_2. Marginal damage in the baseline runs depends on how the higher CO_2 stock reduces the continuation value. When there is a threshold, marginal damage also depends on how the additional CO_2 stock affects the continuation value in the post-threshold regime and on how the additional emissions affect the probability of crossing the threshold before the next period. When the decision-maker wants to avoid a certain threshold, these two effects reduce to the cost of adjusting future policy to still avoid the threshold. If future policy cannot be adjusted because additional emissions would trigger the tipping point, then these additional emissions bear the entire cost of the regime shift. Therefore, when the threshold location is known, its effect on policy depends on the cost of imposing a temperature constraint relative to the cost of switching regimes. If the threshold is above DICE's baseline peak, then the tipping point is irrelevant because the temperature constraint would be slack. If the threshold is too close to the initial temperature, then the cost of avoiding it is extremely large and marginal damage includes the effect of additional CO_2 in the post-threshold regime. Finally, if the threshold is in an intermediate region, then the policy path has temperature approach the

[17] We see a similar effect on the near-term social cost of carbon with the decay rate tipping point because of the long time horizons over which its effects manifest themselves.

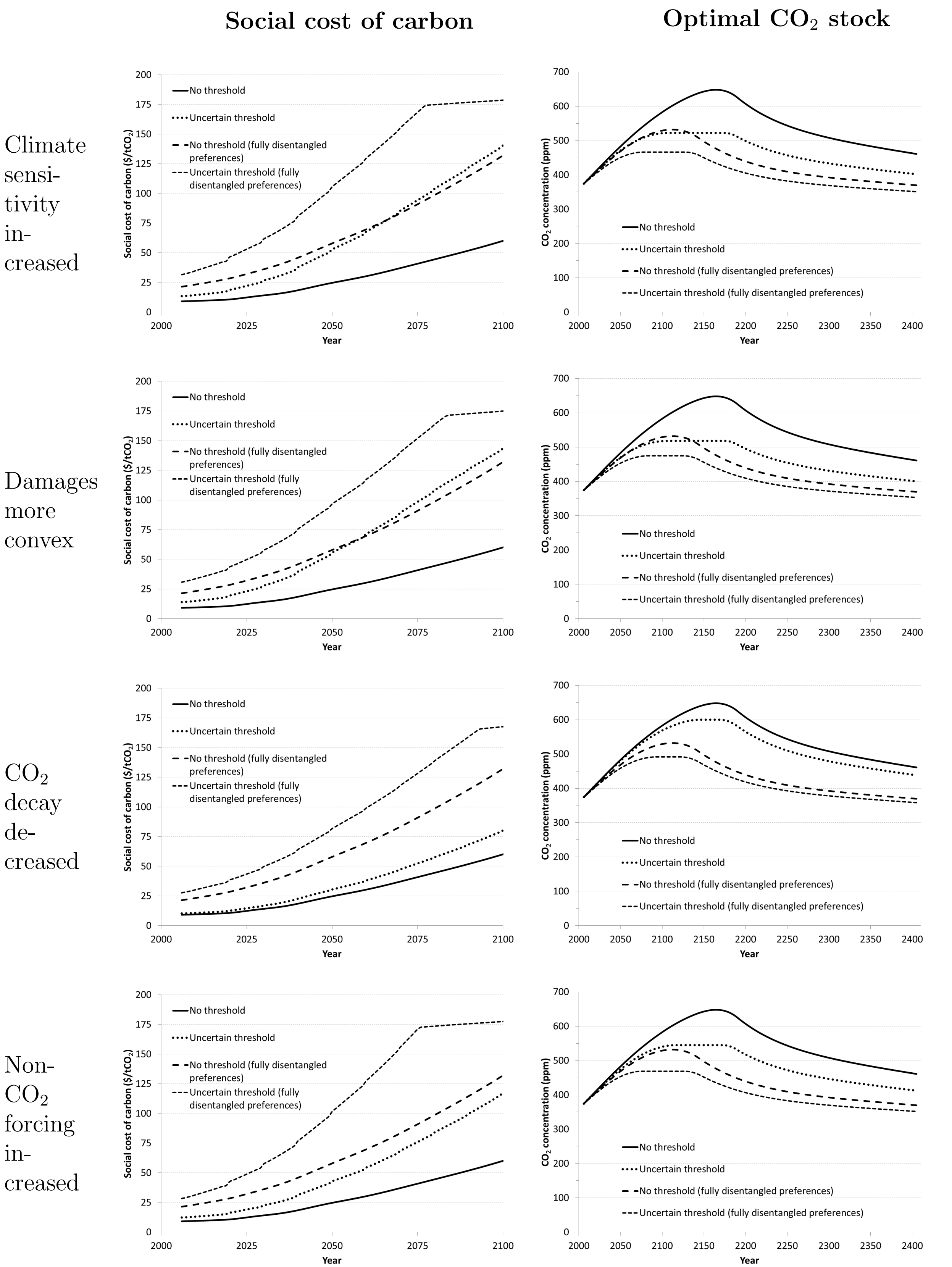


Figure 3.3: Time paths for the social cost of carbon (current value) and the CO_2 stock under each type of tipping point with standard preferences and with fully disentangled preferences. With an uncertain threshold, we simulate a path that happens to never cross a threshold in order to see how the modeled policymaker adjusts to the risk over time.

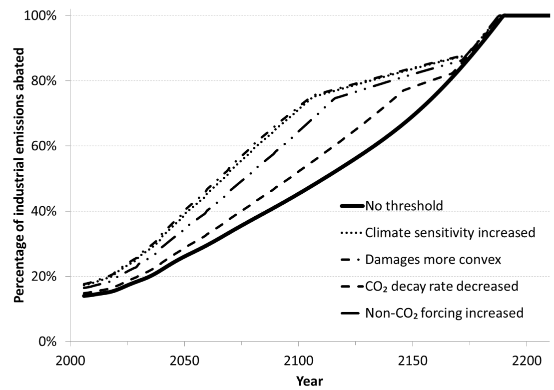


Figure 3.4: The evolution of abatement under expected draws with partially disentangled preferences when the temperature threshold is uncertain and happens to never be crossed.

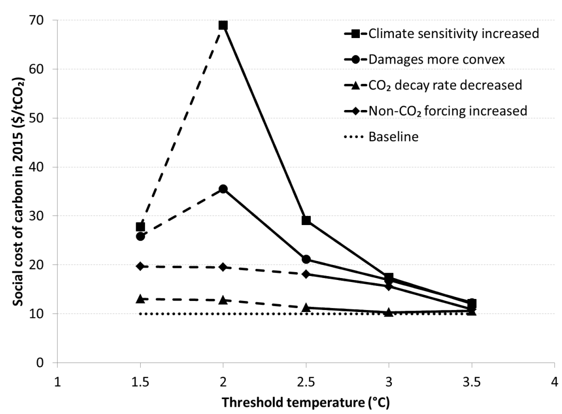
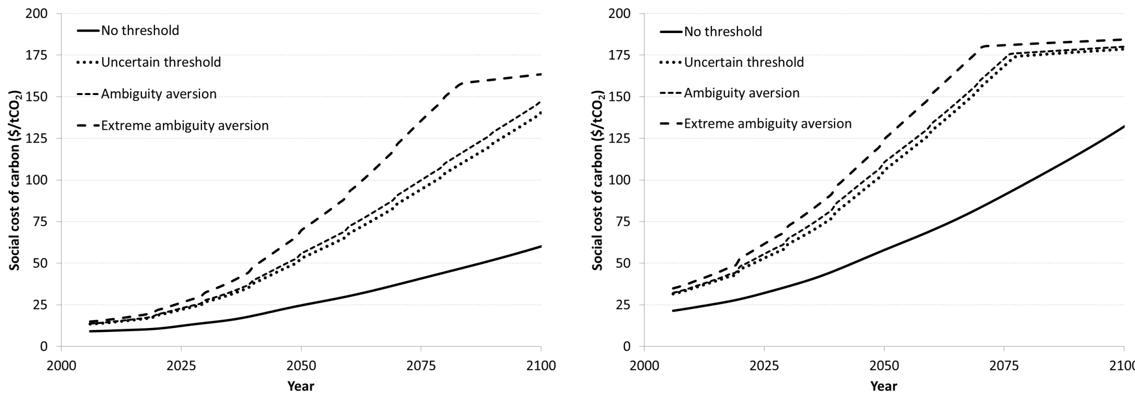


Figure 3.5: The social cost of carbon in 2015 (current value) when the temperature threshold is known to be at a certain level. Solid lines connect points for which the policymaker plans to avoid the threshold. All of these runs use partially disentangled preferences. This figure should only be taken as illustrative for now, as we are still obtaining better approximations for many of these cases.

Table 3.4: The probability that the threshold will ever be crossed. In parentheses, the expected year of crossing, conditional on the threshold being crossed at some point. This will tend to occur earlier when policy makes temperature peak earlier.

	Tipping point			
	Climate sensitivity increased	Damage convexity increased	CO_2 sinks weakened	Non-CO_2 forcing increased
Partially disentangled preferences				
Uncertain threshold	0.47 (2041)	0.46 (2041)	0.59 (2050)	0.51 (2043)
Ambiguity aversion	0.46 (2040)	0.46 (2040)	0.59 (2050)	0.50 (2043)
Extreme ambiguity aversion	0.42 (2037)	0.41 (2037)	0.59 (2050)	0.47 (2041)
Fully disentangled preferences				
Uncertain threshold	0.36 (2035)	0.38 (2036)	0.42 (2038)	0.37 (2035)
Ambiguity aversion	0.36 (2035)	0.38 (2036)	0.41 (2038)	0.36 (2035)
Extreme ambiguity aversion	0.33 (2033)	0.36 (2035)	0.41 (2038)	0.34 (2033)

The baseline policy path has a 66% chance of ever crossing the unknown threshold with standard preferences. It crosses a 2.5°C threshold in 2081, and its expected year of crossing conditional on crossing at some point is 2057. With fully disentangled preferences, the baseline policy path produces a 49% chance of ever crossing the threshold and never crosses 2.5°C. Its expected year of crossing conditional on crossing at some point is 2043. The decision-maker follows the baseline path if unaware of tipping point possibilities.



(a) Partially disentangled preferences

(b) Fully disentangled preferences

Figure 3.6: The effect of ambiguity aversion on the social cost of carbon (current value) in the scenario where the tipping point increases climate sensitivity. Ambiguity aversion does not have any stronger of an effect with the other two tipping points. Temperature follows expected draws, and the threshold is never crossed in these simulations.

Table 3.5: The social cost of carbon ($\$/tCO_2$, current value) in 2015 under expected draws. It is $\$10/tCO_2$ in the baseline case with standard preferences and $\$25/tCO_2$ in the baseline case with fully disentangled preferences.

	Tipping point			
	Climate sensitivity increased	Damage convexity increased	CO_2 sinks weakened	Non-CO_2 forcing increased
Partially disentangled preferences				
Uncertain threshold	16	17	11	14
Ambiguity aversion	16	17	11	14
Extreme ambiguity aversion	18	19	12	15
Fully disentangled preferences				
Uncertain threshold	39	38	34[a]	36
Ambiguity aversion	40	39	33[a]	36
Extreme ambiguity aversion	44	42	35	38

[a] These two values actually differ by only 0.2, suggesting that the apparent decrease in the social cost of carbon with greater ambiguity aversion is probably an artifact of the numerical approximation.

threshold but never cross it.

In contrast, if the threshold is uncertain, then additional emissions at the temperatures reached by DICE never bear the entire cost of a regime shift but instead bear the cost of making the regime shift more likely to occur. With an uncertain threshold, an important part of marginal damage is the effect of additional emissions on the chance of crossing a threshold immediately after the current period. This additional chance is given by the slope of the hazard rate (Figure 3.2b), which is greater when CO_2 concentrations are higher. The effect of additional emissions on the hazard rate therefore raises the social cost of carbon by an amount that increases over time. The increase in the hazard rate is multiplied by the total loss from crossing the threshold to give the expected cost from the one-time increase in threshold risk. Ambiguity aversion primarily affects this total loss from crossing the threshold, so it too matters more with higher CO_2 stocks because of their steeper hazard rate. The effect of both an uncertain threshold and ambiguity aversion grow with time, at least until expected temperatures peak and the hazard rate goes to zero.

As important as thresholds can be, the change from partially disentangled (or standard) preferences to fully disentangled preferences has a greater effect on the social cost of carbon. This is because one of the main near-term tradeoffs in DICE is how much current consumption to sacrifice for the sake of reducing climate damages in a richer future. When the decision-maker is more sensitive to consumption fluctuations over time, sacrificing today becomes less appealing. Separating time and risk preferences (as in the runs with fully disentangled preferences) allows us to reflect recent advances in modeling decision-makers as more averse to consumption fluctuations due to risk than to consumption fluctuations due to time. Representing preferences in this form acts like using a lower discount rate, and DICE and other IAMs are especially sensitive to the choice of discount rate because they include stark intertemporal tradeoffs over long timescales (e.g., Newbold and Daigneault, 2009; Gerst et al., 2010). We have shown that a more general welfare specification calibrated to market returns can act like a much lower discount rate than the market-based one used in DICE.

3.5 Discussion

We have described an original extension of DICE to include the endogenous possibility of climatic tipping points, learning about the temperature threshold that triggers tipping points, and aversion to ambiguity about the threshold's location. We find that optimal policy is more sensitive to tipping points when they have a greater effect on the damages from additional emissions and, in some cases, when additional CO_2 emissions more precisely control whether tipping points occur. Tipping points that increase the quantity of greenhouse gases are less important than tipping points that increase the damage per unit of atmospheric CO_2. More specifically, tipping points that increase the temperature change caused by atmospheric CO_2 or that increase the effects of temperature change on economic output are more important for policy than are tipping points that increase non-controllable emissions or that increase the time for which CO_2 emissions affect atmospheric concentrations. Preference specifications matter greatly when they affect tradeoffs over time, but the ambiguity attitude does not become crucial to our evaluation of tipping point uncertainty until the probability of tipping points becomes greater than in our model's early years. The type of tipping point faced, the precision of the distribution for the temperature threshold, and preferences for intertemporal substitution are more important for near-term policy than is additional aversion to tipping point uncertainty.

We have modeled an irreversible tipping point as occurring upon crossing a threshold in expected temperature, but there are other plausible specifications. First, it is plausible that a threshold's effects only occur after some lag. The extent to which this affects the model depends on the decision-maker's ability to partially or completely "stop" the tipping point once the threshold has been crossed. If the decision-maker only learns that a threshold has been crossed once the system dynamics irreversibly shift, or if the decision-maker learns immediately when a threshold has been crossed but cannot do anything about its future effect on system dynamics, then the primary effect of the lag is merely to decrease the cost of the regime shift by making impacts occur farther in the future. However, if the decision-maker can tell when a threshold has been crossed while still being able to forestall irreversible changes, then introducing a lag should make the decision-maker less cautious about avoiding the threshold but also more active in response to the threshold's crossing.

A second plausible extension is to make the threshold location stochastic.[18] In this case, stabilizing temperature may not remove all threshold risk and learning occurs more slowly. This would lead the decision-maker to choose a lower CO_2 path in order to avoid crossing the threshold by accident. However, the possibility that a threshold could randomly occur even after temperature has stabilized also reduces the value of stabilization. In sum, we expect that making the threshold stochastic would lead the decision-maker to choose a lower CO_2 trajectory over the near- to medium-term but take longer to reach the point where CO_2 concentrations begin to fall.

The policy implications of increasing uncertainty about the threshold location ultimately depend on where the threshold is expected to be. If the certain threshold were so high as to be irrelevant, then uncertainty about its location could make it more relevant. Increasing the spread of the threshold distribution would then raise the social cost of carbon. If, in contrast, the threshold were known to be so low that it would be extremely costly to avoid crossing it, then increasing the spread of the distribution would increase the chance that the threshold is in a region that the decision-maker could more feasibly avoid. However, if the threshold is known to be high enough to make it worth avoiding and low enough to require effort to avoid it, then increasing the spread of the distribution could actually decrease the social cost of carbon by decreasing the effect of additional CO_2 emissions on the chance

[18]This could also correspond to making the threshold depend on temperature realizations rather than on expected temperature.

that a tipping point occurs. The effect of increased uncertainty on the social cost of carbon probably therefore depends on where the threshold is expected to be and on the magnitude of the increase in uncertainty (compare Brozović and Schlenker, 2011). Further, if there is also uncertainty about the number of thresholds, then crossing a first threshold could make later ones more likely by raising future temperatures. The anticipation of this effect should raise the social cost of carbon (and decrease the CO_2 path) prior to crossing the first threshold. Avoiding the first threshold could be an important step to avoiding later ones.

Our conclusions have implications for climate science, economic modeling, and climate policy. First, it is important for IAMs that climate science improve knowledge about both the effects of tipping points on system dynamics and the types of temperature paths that trigger them (Alley et al., 2002, 2003). It is also important to translate tipping point results into the reduced climate models used by IAMs. Which variables might a tipping point affect? Is a tipping point triggered by medium-term average temperature, by short-term temperatures, by interannual variability, or by the rate of warming? What does the distribution for its occurrence look like? How might we expect to learn about tipping point risks? IAMs' conclusions are sensitive to each of these answers and should be updated as the climate science literature progresses.

Second, economic modeling must be clear about which simplifications are likely to be crucial for the results used in policy assessments. We have shown that optimal policy paths are sensitive to assumptions about damage convexity and climate sensitivity, to assumptions about the possibility of tipping points, and to assumptions about whether agents treat time and risk similarly. Modeling exercises that do not vary key parameters or that assume smooth, predictable system dynamics need to be explicit about the omitted factors that tend to push their estimates in a certain direction. This is especially important when all models tend to omit the same factors. In that case, the spread of models' estimates for the social cost of carbon should not directly give the distribution used in policy analysis. Past compilations of IAMs' estimates (e.g., Tol, 2008) described a set of models that omitted climatic features that we have shown could strongly affect the reported results. Further, by building uncertainty into the decision environment, we have shown that the information structure around tipping points has policy consequences. As a result, it is important that other IAMs include uncertainty in a realistic way and vary the information structure. In some cases, certainty makes a variable less relevant, but in other cases, certainty means more effort will be expended to control a variable's effects. Because not much is known about tipping points or how to model them, they constitute an important form of model uncertainty that covers the effects of tipping points as well as knowledge about them.

Third, regarding climate policy, we find that including tipping points can increase DICE's estimate of the year 2015 social cost of carbon by 50% and can decrease DICE's year 2015 industrial CO_2 emissions by over 1 Gt CO_2. Using a recent bottom-up abatement cost curve, this increase in the social cost of carbon could increase the economical emission reductions in the U.S. by 0.25 Gt CO_2 per year (Creyts et al., 2007). Chosen values for the social cost of carbon could play an increasingly important role in the evaluation of government policies (e.g., Interagency Working Group on Social Cost of Carbon, 2010; Masur and Posner, 2010; Greenstone et al., 2011) and could affect the carbon price eventually targeted by carbon taxes and cap-and-trade policies. The best estimate of the social cost of carbon probably does not treat tipping points as impossible, but it is not clear which type of model comes closest to representing the world we face. In any case, much work remains to make IAMs' representation of tipping points more realistic (Hall and Behl, 2006). The challenge when choosing values for the social cost of carbon is one also faced in, among others, the choices of climate projections (Knutti et al., 2010; Lemoine, 2010a), fuel carbon intensity (Plevin, 2010), and interest rates in monetary policy (Hansen and Sargent, 2001): policymakers must consider and combine the results not just of different models but also of different possible structures for a given model. The decision is further complicated when only a small set of model structures has been explored. It is therefore desirable that policy consider both how a given model's predictions change with its assumptions and how to adapt to the results of future modeling efforts. While we have shown how the possibility of a threshold affects the shadow value of CO_2 emissions in a standard IAM, climate policy constitutes buying insurance against not just the assumed possibility of a tipping point but also against the possibility that the available models do not adequately capture future changes in the climate and the economy.

3.6 Appendix: Model calibration

This appendix describes the feedback representation of temperature change, the decay of atmospheric CO_2 over time, and the model's calibration to DICE-2007. We simplify the carbon cycle and temperature change representations from DICE in order to include the tipping point possibilities that might produce a more realistic model.

DICE determines time t surface temperature from the stock of CO_2, from temperature in the previous period, and from the difference in the previous period between temperature at the surface and in the deep ocean.[19] Crost and Traeger (2010) include temperature as a state variable, but because the inclusion of tipping points increases the time required to solve the model, we prefer to avoid the additional state variable. We therefore use a different relationship to capture the influence of time t CO_2 on time t temperature, and we calibrate this representation so as to capture the marginal relationships important for economic evaluation. We model time t temperature change T_t relative to pre-industrial levels as determined by the CO_2 stock M_t and by the net feedback $f_{atm} + f_t$ (Roe, 2009):

$$T_t = \frac{\lambda[R(M_t) + EF_t]}{1 - (f_{atm} + f_t)} = \frac{\lambda[5.35 \ln(M_t/M_{pre}) + EF_t]}{1 - (f_{atm} + f_t)} \ . \tag{3.13}$$

[19]See Hall and Behl (2006) and Warren et al. (2010) for more on DICE's representation of temperature and the carbon cycle.

The function $R(M_t) = 5.35 \ln(M_t/M_{pre})$ gives the additional radiative forcing in W m^{-2} caused by changing CO_2 concentrations from the pre-industrial level M_{pre} to the time t level M_t (Ramaswamy et al., 2001: Table 6.2), and EF_t is the exogenous non-CO_2 forcing in W m^{-2}. The parameter $\lambda = 0.315$ °C (W m^{-2})$^{-1}$ gives the reference system (black body) temperature change per unit of radiative forcing (Soden et al., 2008; Roe, 2009). The sum $f_{atm} + f_t$ is non-dimensional and must be less than 1.

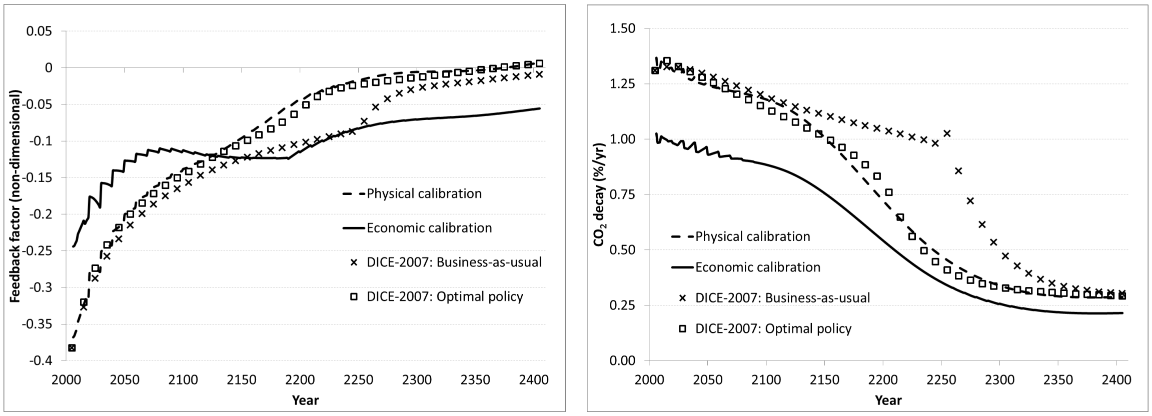
Feedbacks determine the change in temperature generated as the earth system responds to a unit of temperature change in the reference system. When feedbacks are positive, each non-dimensional feedback factor f represents the portion of the total system's temperature change produced by its associated processes. When the CO_2 concentration increases, the atmosphere traps more outgoing radiation (given by $R(M_t)$) even as incoming radiation has not changed. The planet heats up to restore the balance between outgoing and incoming radiation at the top of the atmosphere. This effect is given by the constant λ. However, the increase in surface temperature causes changes in the earth system that in turn cause further changes in surface temperature. For instance, the warmer atmosphere now holds more water vapor, which traps additional outgoing radiation and causes the surface to warm faster. This amplifying response is captured by a positive feedback factor. We assume that feedbacks are independent, linear functions of temperature, which allows us to aggregate the "atmospheric" feedbacks of sea ice, clouds, and water vapor-lapse rate in the constant f_{atm} (Soden et al., 2008; Lemoine, 2010a). Climate sensitivity s is the equilibrium temperature change from doubled CO_2 concentrations:

$$s = \frac{5.35\lambda \ln 2}{1 - f_{atm}} \,, \tag{3.14}$$

where being in equilibrium means $f_t = 0$. DICE-2007 uses a climate sensitivity of 3°C, which implies $f_{atm} \approx 0.61$. We model temperature uncertainty as multiplying s by a lognormally-distributed shock ϵ_t that is independently and identically distributed over time. We implement this by using a quadrature rule for the lognormal product $\epsilon_t s$ and calculating the value of f_{atm} implied by each resulting node.

The time-varying feedback factor f_t represents transient feedbacks. It adjusts equilibrium feedback strength to give time t temperature. DICE-2007 has one state variable for surface temperature and another for deep ocean temperature. The interaction between the two allows the ocean's heat capacity to moderate each period's temperature change. Further, DICE-2007 also delays the effect of radiative forcing on temperature. We use f_t as a reduced-form version of the difference between time t temperature and equilibrium temperature for a given CO_2 concentration (Baker and Roe, 2009). When CO_2 concentrations are increasing, f_t should always be a negative feedback because ocean heat uptake prevents all of the equilibrium surface warming from occurring immediately. When CO_2 concentrations are decreasing, f_t can be positive as the ocean transfers stored heat to the atmosphere.

We calculate the transient feedback f_t as a function of time and the CO_2 stock: $f_t \equiv f(M, t)$. The negative feedback due to ocean heat uptake should weaken (i.e., move towards



(a) Transient feedback (b) CO_2 decay

Figure 3.7: The transient feedback and CO_2 decay rate along the optimized paths for the physical calibration (f_t and ϕ_t) and for the economic calibration (f_t' and ϕ_t'). Also shows the values inferred from DICE-2007 along the optimized policy path (f_{opt} and ϕ_{opt}) and the business-as-usual path (f_{bau} and ϕ_{bau}).

0) with time when the CO_2 stock is constant (Baker and Roe, 2009). The atmospheric CO_2 stock proxies for different emission paths, which means that, all else equal, a higher atmospheric CO_2 stock at a given time might produce a lower (i.e., more negative) f_t as the ocean and atmosphere have had less chance to equilibrate. We define the function $f(M,t)$ in three steps. First, we calculate the f_t implied along the business-as-usual (no policy) and optimized runs in DICE-2007 using equation (3.13). Second, we approximate the function along the time dimension for each of these DICE-2007 runs using piecewise third-order polynomial splines. This gives two approximated functions: $f_{bau}(t)$ and $f_{opt}(t)$. Third, for any (M,t) node for which we want to calculate f_t in our model, we determine $f(M,t)$ by weighting the values $f_{bau}(t)$ and $f_{opt}(t)$ according to which DICE model run has a time t CO_2 value closer to M:

$$f(M,t) = \begin{cases} \max\left\{ \frac{f_{bau}(t)-f_{opt}(t)}{M_{bau}(t)-M_{opt}(t)}[M - M_{opt}(t)] + f_{opt}(t),\ f_{bau}(1) \right\} & \text{if } M_{bau}(t) \neq M_{opt}(t) \\ f_{bau}(t) & \text{if } M > M_{bau}(t) = M_{opt}(t) \\ f_{opt}(t) & \text{if } M \leq M_{opt}(t) = M_{bau}(t) \end{cases}, \tag{3.15}$$

where $M_{bau}(t)$ and $M_{opt}(t)$ are the time paths of the CO_2 stock in the business-as-usual and optimized DICE-2007 runs. The transient feedback at time t is a linear extrapolation from $f_{bau}(t)$ and $f_{opt}(t)$, provided it is no lower than its initial value in the DICE business-as-usual run. Figure 3.7a plots $f_{bau}(t)$ and $f_{opt}(t)$. It also plots f_t along the optimized (M,t) path from our base model without either temperature uncertainty or a threshold. These values are at least broadly similar to the results of Baker and Roe (2009).

We also simplify the representation from DICE-2007 of the evolution of atmospheric CO_2. DICE-2007 has one state variable for atmospheric carbon, another for shallow ocean carbon, and a third for deep ocean carbon. These state variables and their associated transition matrix constitute a simple carbon cycle model determining how atmospheric CO_2 decays from period to period. We instead have an explicit decay rate ϕ_t that is a function of time t and of the atmospheric CO_2 stock M_t. As with the transient feedback factor f_t, including the atmospheric CO_2 stock as an argument proxies for the emission path up to time t. The decay rate ϕ implied by a change in DICE carbon stocks from M_t to M_{t+10} (where t is in years) solves the following equation:

$$M_{t+10} = M_{pre} + (1-\phi)^{10}(M_t - M_{pre}) + E_t \frac{1-(1-\phi)^{10}}{\delta} \ , \tag{3.16}$$

where E_t gives annual CO_2 emissions over the decade and M_{pre} is the pre-industrial CO_2 stock. The algorithm for approximating the function $\phi_t \equiv \phi(M,t)$ is essentially as described for f_t, except using equation (3.16) in place of equation (3.13). However, because the decay rate may be controlled more by time than by the emission path, we do not extend the linear approximation beyond $M_{bau}(t)$ and $M_{opt}(t)$. We instead use the following rule that linearly interpolates $\phi(M,t)$ only when M is between $M_{bau}(t)$ and $M_{opt}(t)$:

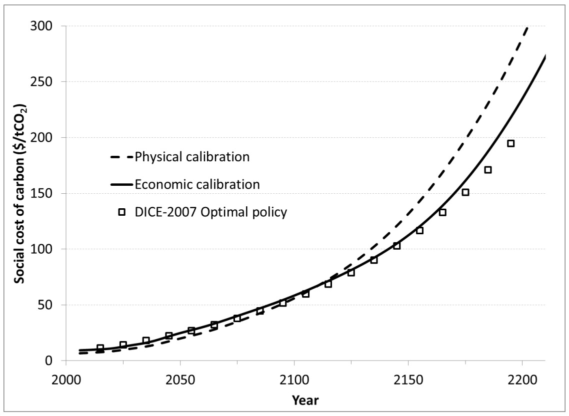
$$\phi(M,t) = \begin{cases} \phi_{bau}(t) & \text{if } M > M_{bau}(t) \\ \left(1 - \frac{M - M_{opt}(t)}{M_{bau}(t) - M_{opt}(t)}\right)\phi_{opt}(t) + \frac{M - M_{opt}(t)}{M_{bau}(t) - M_{opt}(t)}\phi_{bau}(t) & \text{if } M_{opt}(t) < M \leq M_{bau}(t) \\ \phi_{opt}(t) & \text{if } M \leq M_{opt}(t) \ , \end{cases} \tag{3.17}$$

Figure 3.7b shows $\phi_{bau}(t)$, $\phi_{opt}(t)$, and, along the optimized path, ϕ_t.

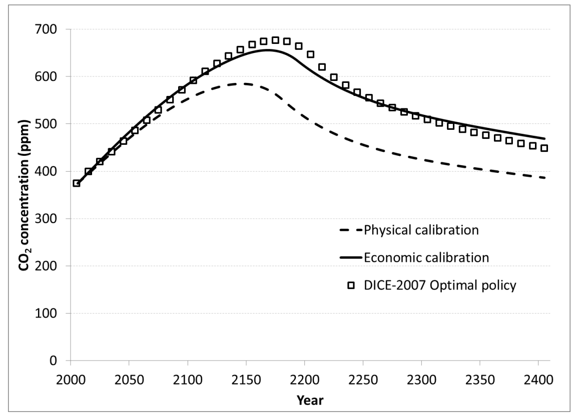
These inferred parameters f_t and ϕ_t enable us to replicate the CO_2-temperature relationship from DICE-2007 as well as the relation between CO_2 levels with decadal timesteps.[20] We label this parameterization the "physical calibration". However, these two inferred parameters do not reproduce the dynamics of the CO_2 stock and the social cost of carbon in DICE (Figure 3.8). Because we are primarily interested in how these policy-relevant values change under different specifications, we use an "economic calibration" that better matches DICE-2007's optimized social cost of carbon.

First, our use of an annual timestep instead of a decadal timestep tends to make the inferred CO_2 decay rate different than what we would need to replicate CO_2 dynamics with emissions varying from year to year. In particular, emissions in the second century often decline over a decade while DICE's stock transition equations treat annual emissions at the start of the decade as lasting for all ten years. We therefore adjust the CO_2 decay rate as

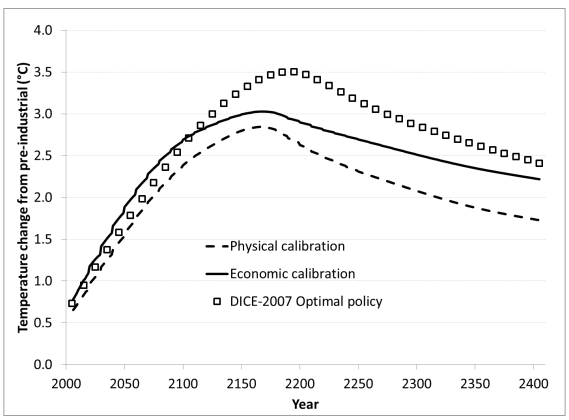
[20] We could further refine the calibration by obtaining DICE output along lower emission paths in order to provide data points for the decay rate and the transient feedback that are closer to the emission paths we see with the threshold possibility. Note, however, that DICE's own carbon cycle model was not calibrated to emission paths like those seen in optimized scenarios.



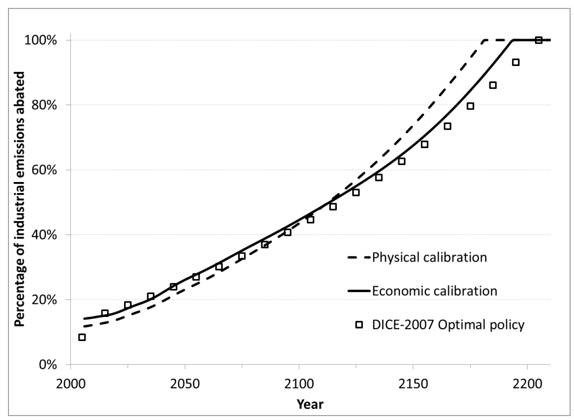
(a) Social cost of carbon (current value)



(b) CO_2 concentration



(c) Temperature relative to the pre-industrial level



(d) Abatement

Figure 3.8: The optimized paths for the physical calibration, for the economic calibration, and for DICE-2007. The plot for the social cost of carbon shows a shorter time series because DICE calculates the social cost of carbon as the cost of the next unit of abatement. Once DICE reaches full abatement around the year 2200, the cost of the next unit of abatement is lower than the marginal damage. Our model can calculate the social cost of carbon from marginal damage or from the marginal abatement cost.

follows in the economic calibration (Figure 3.7b):

$$\phi_t' = 0.75\phi_t \ . \tag{3.18}$$

This experimentally determined adjustment enables us to better match DICE's CO_2 stock dynamics.

Second, while the transient feedback in the physical calibration reproduces the relationship between CO_2 stocks and temperature over representative time paths, the physical calibration does not reproduce the social cost of carbon because it alters the marginal relationship between CO_2 and temperature by failing to delay the effect of radiative forcing on temperature. In Figure 3.8, the generally greater social cost of carbon in the physical calibration leads to greater abatement, which reduces both the CO_2 stock and temperature. To better capture the social cost of carbon's time path, the economic calibration uses the following experimentally determined transient feedback f_t' (Figure 3.7a):

$$f_t' = f_t + \max\{-0.06,\, 0.125 - 0.001t\} \ , \tag{3.19}$$

where $t = 0$ corresponds to 2005. This adjustment makes the transient feedback less negative (increasing temperature for a given CO_2 stock) up to the year 2130, and after that time it makes the transient feedback more negative (decreasing temperature for a given CO_2 stock). As Figure 3.8 shows, the economic calibration largely reproduces the time paths for the social cost of carbon, abatement, and the CO_2 stock, but this comes at the cost of making temperature 0.5°C too low near its its peak. The economic calibration better represents the marginal effect of CO_2 emissions in DICE by reducing the effect of the total CO_2 stock on temperature, which increases available economic output relative to DICE.

3.7 Appendix: Model specification

This appendix provides the transition equations for state variables and exogenous variables (see Nordhaus, 2008; Crost and Traeger, 2010). It also describes the four modeled tipping points in terms of these equations and illustrates how post-threshold policy differs from policy in the absence of tipping points.

The transition equations for the state variables of effective capital (k_t) and atmospheric CO_2 (M_t) are:

$$k_{t+1} = e^{-(g_{L,t}+g_{A,t})}\left[(1-\delta_k)k_t + (1-\Psi_t\mu_t^{a_2})\frac{Y_{gross}}{1+D} - c_t\right] \tag{Capital}$$

$$M_{t+1} = M_{pre} + (1-\phi_t')(M_t - M_{pre}) + \sigma_t(1-\mu_t)Y_{gross} + B_t \ . \tag{CO$_2$}$$

In the transition equation for CO_2, σ_t is the emission intensity of gross output and B_t gives exogenous CO_2 emissions from non-industrial sources such as land use change. M_{pre} is the pre-industrial CO_2 stock, and ϕ_t' is calculated as detailed in the first appendix. The first term in the capital transition equation has capital depreciating at constant rate δ_k, and the last two terms define capital investment as any available output not allocated to the control variables of consumption c_t and abatement. Here, Ψ_t and a_2 give the cost of abating the chosen fraction μ_t of emissions. The term outside the brackets adjusts for the growth of labor and technology to keep capital in effective terms. Gross output Y_{gross} is a function of the capital stock:

$$Y_{gross} = k_t^{\kappa} \ . \qquad \text{(Gross output)}$$

The parameter κ gives the capital elasticity in a Cobb-Douglas production function. Climate damages D reduce gross output more strongly as temperature increases:

$$D = b_2 T_t^{b_3} \ , \qquad \text{(Damages)}$$

where $b_3 = 2$ in DICE-2007 and temperature change T_t relative to pre-industrial levels is as in the first appendix.

The transition equations for the exogenous variables are as in DICE-2007, but adjusted for the annual timestep. We here list them and give the parameterization in Table 3.6. In each case, $t = 0$ corresponds to the year 2005. See Crost and Traeger (2010) for more details on variable definitions and implementation.

$$\begin{aligned}
A_t =& A_0 \exp\left[\frac{g_{A,0}}{\delta_A}\left(1 - e^{-t\delta_A}\right)\right] && \text{(Production technology)}\\
g_{A,t} =& g_{A,0} e^{-t\delta_A} && \text{(Growth rate of production technology)}\\
L_t =& L_0 + (L_\infty - L_0)\left(1 - e^{-t\delta_L}\right) && \text{(Labor)}\\
g_{L,t} =& \delta_L \left[\frac{L_\infty}{L_\infty - L_0} e^{t\delta_L} - 1\right]^{-1} && \text{(Growth rate of labor)}\\
\beta_t =& \exp\left(-\rho + (1-\eta) g_{A,t} + g_{L,t}\right) && \text{(Effective discount factor)}\\
\sigma_t =& \sigma_0 \exp\left[\frac{g_{\sigma,0}}{\delta_\sigma}\left(1 - e^{-t\delta_\sigma}\right)\right] && \text{(Uncontrolled emissions per output)}\\
\Psi_t =& \frac{a_0 \sigma_t}{a_2}\left(1 - \frac{1 - e^{t g_\Psi}}{a_1}\right) && \text{(Abatement cost factor)}\\
B_t =& B_0 e^{t g_B} && \text{(Non-industrial } CO_2 \text{ emissions)}\\
EF_t =& EF_0 + 0.01(EF_{100} - EF_0)\min\{t, 100\} && \text{(Non-}CO_2 \text{ forcing)}
\end{aligned}$$

The constraints prevent the decision-maker from using more than the output available after accounting for damages and from abating more than 100% of emissions in a period:

$$c_t + \Psi_t \mu_t^{a_2} \leq \frac{Y_{gross}}{1 + D} \qquad (3.20)$$

$$\mu_t \leq 1 \ . \qquad (3.21)$$

When the constraint in equation (3.20) is slack, we have positive capital investment, and when the constraint in equation (3.21) is slack, economic activity produces some CO_2 emissions that are not abated.

The challenge in solving the model lies not in finding the optimal actions for a given value function but in determining the value functions that satisfy the relations in equations (3.3) and (3.5) (see Kelly and Kolstad, 1999, 2001). We begin with a guess for the value function and a set of Chebychev nodes in the three-dimensional state space. We then use the initial guess for the continuation value to find each node's optimal controls c_t^* and μ_t^* and optimal value. Knowing the optimal value at each Chebychev node, we approximate the value function across the rest of the state space using a set of Chebychev basis polynomials (Miranda and Fackler, 2002). We repeat the process using this approximated value function as the new initial guess, with iteration continuing until the coefficients of the value approximant's basis functions change by less than 0.0001. When the temperature threshold is uncertain, the pre-threshold value function is smooth over the relevant state space because the hazard rate changes smoothly with time and the CO_2 level. However, when the threshold is known to be at T^*, each node in the state space is associated with either the pre-threshold or the post-threshold value function, and the combined value function exhibits a discontinuity between regions in which optimal pre-threshold policy would and would not avoid it. We undertake numerical approximation in this case by making the threshold stochastic in order to smooth out the value function (compare Brozović and Schlenker, 2011). This stochasticity comes from placing a low-variance normal distribution around T^*, which means that transitioning to a temperature above T^* is virtually–but not totally—certain to cross the threshold. As the variance of this distribution approaches zero, we approach the case with a non-stochastic, certain threshold.[21]

We now describe how each post-threshold regime affects damages D, ignoring transient feedbacks for simplicity. First, the climate sensitivity regime uses $\hat{s} = 2s = 6°\text{C}$, corresponding to $f_{atm} \approx 0.81$ (see the first appendix). This change doubles equilibrium temperature, which increases equilibrium damages from $b_2 T^2$ to $b_2(2T)^2$, where T is calculated using a climate sensitivity of 3°C. Equilibrium damages therefore quadruple, though the presence of transient feedbacks changes the actual effect on time t damages. Second, increasing the convexity of damages means using $\hat{b}_3 = b_3 + 1 = 3$. This multiplies damages by T, which is always less than 4 in expectation in our baseline runs but can be greater under some temperature realizations. Third, weakening CO_2 sinks means using $\hat{\phi}_t = (\phi_t')/4$, which increases the length of time for which a unit of CO_2 emissions affects the atmospheric CO_2 stock. The change from ϕ to ϕ' only has a significant effect on the CO_2 stock once enough time has passed for the additional accumulation to matter. Finally, increasing exogenous non-CO_2 forcing means using $\widehat{EF}_t = EF_t + 1.5$. This increases temperature by 1.2°C in equilibrium,

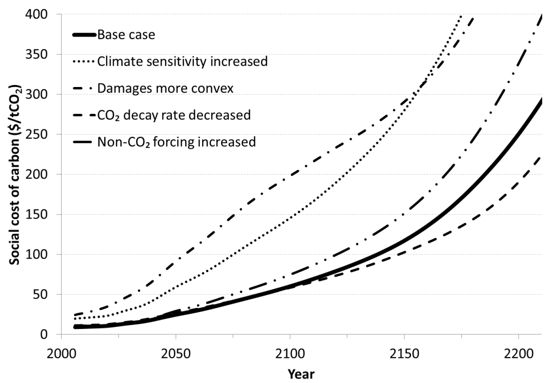
[21] The certain threshold model is currently solved with standard deviations between 0.01°C and 0.40°C, depending on the regime and threshold level under consideration. We are working on better approximations and will eventually compare different standard deviations to assess the effect of uncertainty.

Table 3.6: Parameterization of the transition equations. See Nordhaus (2008) and Crost and Traeger (2010) for more information. Also shows how these parameters change in the post-threshold regimes.

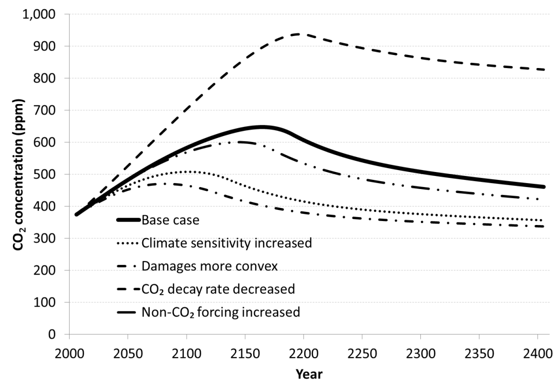
Parameter	Value	Description
A_0	0.02722	Initial production technology
$g_{A,0}$	0.009	Initial annual growth rate of production technology
δ_A	0.001	Annual rate of decline in growth rate of production technology
L_0	6514	Population in 2005 (millions)
L_∞	8600	Asymptotic population (millions)
δ_L	0.035	Annual rate of convergence of population to asymptotic value
σ_0	0.131418	Initial emission intensity before emission reductions (GtC/output)
$g_{\sigma,0}$	-0.00730	Initial annual growth rate of emission intensity
δ_σ	0.003	Annual change in growth rate of emission intensity
a_0	1.17	Cost of backstop technology in 2005 ($1000/tC)
a_1	2	Ratio of initial backstop cost to final backstop cost
a_2	2.8	Abatement cost exponent
g_Ψ	-0.005	Annual growth rate of backstop cost
B_0	1.1	Initial non-industrial CO_2 emissions (GtC/y)
g_B	-0.01	Annual growth rate of non-industrial emissions
EF_0	-0.06	Initial exogenous forcing from non-CO_2 greenhouse gases (W m^{-2})
EF_{100}	0.30	Year 2105 exogenous forcing from non-CO_2 greenhouse gases (W m^{-2})
κ	0.3	Capital elasticity in Cobb-Douglas production function
δ_κ	0.1	Annual depreciation rate of capital
b_2	0.0028388	Coefficient of temperature in the damage function
b_3	2	Exponent on temperature in the damage function
s	3	Climate sensitivity (°C)
M_{pre}	596.4	Pre-industrial atmospheric CO_2 (GtC)
ρ	0.015	Annual rate of pure time preference
k_0	$137/(A_0L_0)$	Initial effective capital, with initial capital stock of 137 US$trillion
M_0	808.9	Initial atmospheric CO_2 (GtC)
Parameters for post-threshold regimes (i.e., for tipping points' effects)		
$\hat{s}$	$2s$	Climate sensitivity increased
$\hat{b}_3$	b_3+1	Damages more convex
$\hat{\phi}_t$	$(\phi'_t)/4$	CO_2 sinks weakened
$\widehat{EF}_t$	$EF_t+1.5$	Non-CO_2 forcing increased

which means damages increase in equilibrium to $b_2(T+1.2)^2 = b_2(T^2+2.4T+(1.2)^2)$. When temperature is greater than 2.3°C, the effect of this tipping point on damages is less than the equilibrium quadrupling or expected tripling caused by the climate sensitivity and damage convexity tipping points. For example, when the temperature is 3°C, the increase in non-CO_2 forcing approximately doubles damages.

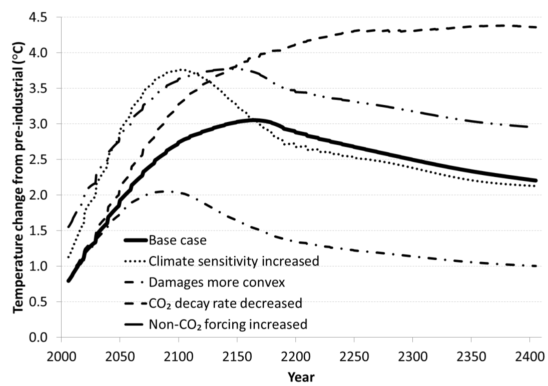
Figure 3.9 shows how optimal policy, temperature, and CO_2 concentrations would evolve if each tipping point occurred exogenously in 2005. Each regime affects policy and the climate differently, which in turn affects the degree to which the decision-maker tries to avoid that type of tipping point. Optimal policy responds in accord with how each tipping point affects damages. Increasing climate sensitivity or increasing damage convexity both raise the social cost of carbon and lower the CO_2 stock path. However, they have quite different effects on temperature because a given CO_2 stock produces higher temperatures when climate sensitivity is increased while a given temperature produces greater damages when the convexity of the damage function is increased. These results indicate how, aside from tipping points, IAMs' results are sensitive to assumptions about the uncertain parameters determining climate sensitivity and the convexity of the damage function. Decreasing the decay rate of CO_2 does not significantly affect abatement or the social cost of carbon, but the reduced stock decay does eventually produce higher CO_2 concentrations, temperatures, and damages. Emission decisions have impacts for a longer time than usual, but that change does not greatly affect the present value of the damage they produce. Finally, while increasing non-CO_2 forcing does increase temperature and damages, it does not affect the CO_2 stock path or the social cost of carbon as strongly as do the climate sensitivity or damage convexity tipping points. Entering this regime reduces economic output, but this effect does not interact as strongly with emission decisions.



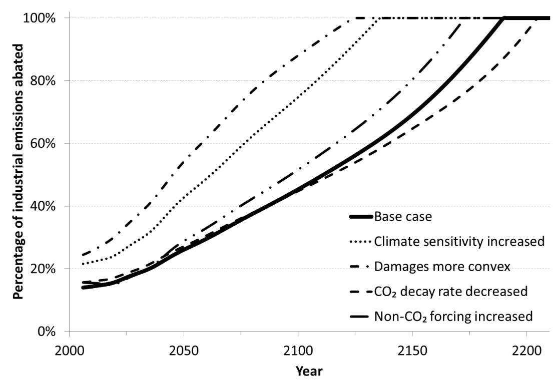
(a) Social cost of carbon (current value)



(b) Optimal CO_2 concentration



(c) Optimal temperature change from pre-industrial level



(d) Abatement

Figure 3.9: The evolution of the optimal social cost of carbon, CO_2 concentration, temperature, and abatement under expected temperature outcomes if each type of tipping point were crossed in 2005. These plots show how optimal policy and the climate respond to each tipping point's occurrence in model runs with partially disentangled preferences.

Chapter 4

Using the magnitude and duration of an emission tax to stimulate adoption of green technology[1]

Pollution policy is often set in anticipation of lower-pollution technology. Firms' adoption of lower-pollution technology can reduce the optimal emission tax, but lowering the tax in anticipation of the technology's adoption reduces the incentive to actually adopt it. Whereas the price of tradable permits automatically responds to firms' investments in low-pollution technology, a linear tax does not. We show that an ex ante emission tax often cannot obtain a socially efficient outcome because it must fix both the incentive to invest and the incentive to reduce emissions. Many outcomes with socially optimal emissions from each type of plant have suboptimal investment, and many outcomes with socially optimal investment result from the regulator using a tax that is set inefficiently high in order to stimulate investment. The numerical multiperiod setting makes adopted technology last for two periods and allows the regulator to control the duration of the tax. Firms recognize that the next period's tax depends on the fraction of firms that invests. The regulator chooses a longer tax when there is little low-pollution plant already in the fleet and firms expect their investments to strongly reduce the next period's tax, but the cost of not being able to adjust the tax limits its use. Investments and emissions often are not socially optimal, and the long-run level of low-pollution plant is not sensitive to the initial level of low-pollution plant.

[1]Many thanks to my co-author, Larry Karp.

4.1 Introduction

Near-term climate policy must do more than constrain emissions: it should also induce firms to invest in reducing the cost of future, more ambitious policy. Desired investments include research and development into low-pollution technologies, installation of carbon capture and storage technology on power plants, build-out of fueling infrastructure for next-generation vehicles, and construction of factories for mass-producing photovoltaic cells and batteries. It is not uncommon for policy to have this dual goal of mitigating pollution while obtaining irreversible investment that reduces the cost of mitigating pollution. With sulfur dioxide, electricity generators can choose the type of plant they build and whether to install scrubbers. Reducing nitrogen-rich runoff from fertilizer application may be not just a matter of using less fertilizer but also of adopting monitoring technology that enables more targeted application. Unless marginal damage is constant, the optimal pollution tax should reflect realized investments, but it may be politically or informationally infeasible for the regulator to condition the tax on realized investment. Then, in contrast to a program of tradable permits, a tax fixes investment and emission incentives independently of other firms' investments. The optimal pollution affects investment even as it is affected by anticipated investment; it is caught between needing to adjust to investments' effect on abatement cost while not undercutting the incentive to invest. The regulator often must choose between a tax that is optimal for some given level of investment and a tax set so as to change the level of investment, but, in many cases, neither will produce a socially optimal outcome.

We develop models that illuminate the interaction between the optimal pollution tax and firms' decisions to adopt "green" technology that reduces the benefit from polluting. The one-period analytic sequential game provides insight as to when the regulator departs from the standard rule of equating marginal damage and marginal abatement cost. Because the tax serves not only to reduce emissions but also to select investment equilibria, the regulator might set the tax high enough to stimulate investment even when such a tax would not be optimal after investment has occurred. Yet, even if a regulator cannot condition a tax on investment within some sufficiently short time horizon, it may have a chance to update the tax over a longer time horizon. If previously installed low-pollution technology is still in use when the regulator moves to update the tax, lowering the rate would undercut the initial incentive to invest. The regulator may therefore want to commit in advance to not change the tax. The duration of the tax becomes an additional decision variable by which it can mitigate the tradeoff between marginal damage pricing and obtaining optimal investment. Our multiperiod numerical model explores this option in a dynamic game between the regulator and a continuum of firms. Here, low-pollution plant lasts for two periods, which makes firms' investment decisions depend on their expectation of the following period's tax policy, and the regulator can choose whether a tax lasts for one or two periods. We assess the factors that drive the choice of a longer or shorter tax and examine the degree to which controlling duration can bring equilibrium investment closer to socially optimal investment.

Our results extend the literature on policy instruments' incentives for the development, adoption, and diffusion of new technology (see Fischer et al., 2003; Requate, 2005). We consider the implications of a price instrument when firms can make lumpy investments in adopting technology that lowers marginal abatement cost. Requate and Unold (2003) described three settings. First, if the regulator does not anticipate that homogeneous firms could adopt low-pollution technology, then an otherwise optimal ex ante tax tends to lead to over-investment by firms because it is set too high. In contrast, an otherwise optimal ex ante cap tends to lead to under-investment as some firms free ride on the lowered permit prices produced by other firms' investments. The additional investment incentive produced by myopic taxes decreases in the presence of abatement cost uncertainty (Zhao, 2003). Second, if the regulator sets the tax or cap after firms have made their investments, then either instrument can achieve socially optimal investment and pollution. Karp (2008) extended this second setting to the case of a global game in which each firm is uncertain about other firms' investment cost. Third, if the regulator commits to the policy before firms make their investment decisions, then using a cap produces the socially optimal outcome because the price of tradable permits responds to firms' investments. In contrast, using a tax produces multiple equilibria because the payoff to investing does not respond to realized investments. An ex post tax acts like the permit price in responding to realized investment, but an ex ante tax does not adjust to realized investment. Weber and Neuhoff (2010) extended this literature to the case of hybrid price-quantity instruments fixed before heterogeneous firms invest in innovation. We consider how a regulator using an ex ante tax and facing firms differentiated only by past adoption of low-pollution technology chooses which equilibrium to play: an equilibrium with investment (i.e., adoption) or an equilibrium without investment. We show that recognizing the endogeneity of investment can raise the regulator's optimal tax above the level that equates marginal abatement cost and marginal damage. In addition, while making investment costs heterogeneous enables equilibria in which only some firms invest, this outcome will still usually not be socially efficient.

When we move to a multiperiod setting, firms must compare the value from adopting low-pollution technology today to the value from having the option to do so in the following period. This structure is that of a real options problem, though without the uncertainty that usually makes optionality of special interest. When the option is to undertake a predefined, discrete emission reduction policy, sunk abatement costs combine with uncertainty to make a social planner more hesitant to adopt the policy (Pindyck, 2000, 2002; Fisher and Narain, 2003). When the policy instrument is continuous, the effects of abatement capital's irreversibility are more complex.[2] Under uncertainty about future production technology, making investment irreversible can reduce both the privately optimal and the socially optimal abatement capital by the same fraction. In this case, excess emissions and, as a result, the optimal policy are greater when capital is reversible (Jou, 2001). van Soest (2005) con-

[2]Related work has also considered the implications for the timing of technology adoption of a combination of uncertainty about future policy and uncertainty about either innovation or the cost of adoption (e.g., Isik, 2004; Dehghani, 2011).

sidered the effect of myopic taxes and non-tradable permits on the timing of adoption of new technologies when they arrive stochastically and the cost of adopting them is irreversible.[3] von Döllen and Requate (2008) considered more informed policies in a setting where a better low-pollution technology arrives at some unknown future time. Investment into either the current or the future low-pollution technology is irreversible. In this case, either taxes or tradable permits can achieve the social optimum if set ex post (after the technology arrives). However, they cannot achieve the social optimum if the regulator sets the policies myopically or if the regulator ex ante determines policy for an initial period and also determines a potentially different policy to take hold after the technology's expected date of arrival. Finally, if the regulator ex ante sets one policy for an initial period and another that will take effect once the new technology actually does arrive (but before firms decide on its adoption), then tradable permits can achieve the social optimum while, as in Requate and Unold (2003), taxes produce multiple equilibria. This last case with taxes is analogous to our model, which additionally considers the regulator's ability to select which equilibrium to play while looking only at the period in which the technology actually arrives.

Our multiperiod setting captures aspects of a hold-up problem. Durable investment in low-pollution technologies may be chilled if firms expect their investments to lead the regulator to lower the tax or permit price in the subsequent period. In our case, the regulator can use a longer tax to mitigate the hold-up problem at the cost of not being able to fine-tune the tax to each period's technological and investment environment.[4] Similar hold-up problems have been considered before. After irreversible (often lumpy) investments have been made, regulators have an incentive to change policies because of multiple objectives (Abrego and Perroni, 2002; Helm et al., 2004) or because of the ability to use permit markets to reduce firms' licensing profits without running afoul of patent law (Laffont and Tirole, 1996). Further, firms might invest strategically before a policy is set so as to affect the optimal stringency of future policy (Malik, 1991; Biglaiser et al., 1995; Gersbach and Glazer, 1999). For example, if firms are large enough for their lumpy investments to affect the regulator's subsequently chosen policy, then firms have an incentive to under-invest in technology that enables emission reductions (Gersbach and Glazer, 1999). However, tradable permits mitigate this problem by allowing firms that invest to not only save on emission costs but also to earn revenue on the permit market. The regulator in this setting can achieve the socially optimal outcome if it commits to using a quantity policy but not if it commits to using a price policy. In contrast, in our setting, investment occurs after the regulator sets the level

[3]van Soest (2005) argued that tradable permits would provide less incentive than taxes to speedily adopt new technologies because the permit price would decrease as adoption increases while the tax would remain the same. Note, however, that their result assumes that both the cap and the tax are set myopically. We show that an ex ante tax set with awareness of possible new technologies would in general offer less incentive to invest than would a cap because the tax responds to anticipated investment before the investment is realized.

[4]When our regulator announces a two-period tax, it is locked into that policy and firms therefore view the announcement as credible.

of the policy and firms are not large enough to affect policy on their own.

One way to mitigate this hold-up problem is to adjust the mechanism for updating policy or the length of commitment periods (e.g., Helm et al., 2003; Fankhauser and Hepburn, 2010). Maintaining policy flexibility has value when technology and the environment are changing or when the regulator anticipates that its information will change, but flexibility also has several costs, including two associated with firms' options to invest in abatement capital. First, flexibility can increase the ability of firms to manipulate future policies via strategic investments. This is a version of the firms holding up the regulator. Greater uncertainty about damages then favors a policy that gives the regulator future discretion while less uncertainty favors using rules to set future policy (Malik, 1991; Tarui and Polasky, 2005). Second, greater flexibility can undercut non-strategic firms' incentives to make long-lived investments, as these investments simultaneously affect and depend on the stability of policy over time (Blyth et al., 2009). The regulator might take advantage of firms having made such investments, or the regulator might learn about the damage function and want to act on that information. Our multiperiod setting assesses the tradeoff between the flexibility to adapt policy ex post to sunk investments and the need to provide ex ante incentives to make long-term investments. The regulator chooses whether the emission tax lasts for one period or for two, with the benefit of a one-period policy being that the regulator can adjust in the next period to the first period's investments but the cost being that firms in the first period will anticipate this adjustment. If the adjustment involves lowering taxes to reflect reduced marginal abatement cost, then firms in the first period have less incentive to make investments that last for more than one period.

Instrument choice is sensitive to the mechanism for updating policy (Requate, 2005; Hepburn, 2006; Montgomery and Smith, 2007). When firms invest strategically and marginal damage is constant, an emission tax obtains the first-best pollution and investment outcomes because the optimal tax does not depend on investment and so firms' strategic ability does not matter. In contrast, a quantity instrument is undercut by the regulator's time inconsistency (Biglaiser et al., 1995). While a linear damage function may be justified for a slow-changing global stock pollutant like carbon dioxide, many flow pollutants are more localized and will have a nonlinear damage function over the policy-relevant range.[5] We model this case of flow pollutants with a strictly convex damage function, which makes the optimal tax depend on firms' investment decisions. Our firms are too small to affect optimal policy by their own actions, but their aggregate decisions do affect optimal policy. We consider how lengthening commitment periods can reduce the inefficiency of an emission tax stemming from the combination of a nonlinear damage function and firms' ability to anticipate future regulation.

We next introduce the one-period analytic model, which describes the possible equilibria and the conditions under which the regulator stimulates investment with a high tax on

[5] When assessed at the level of global policy set for longer timescales, carbon dioxide may also have a nonlinear damage function.

emissions. It shows that equilibrium investment is often not socially optimal. Further, when equilibrium investment is socially optimal, the chosen tax is often higher than socially optimal. We also show that introducing heterogeneity to investment costs avoids the multiple equilibria complication but generally does not provide social optimality. We then describe the numerical model and give results regarding equilibrium investment, tax rates, and policy duration. We conclude with next steps and implications for policy.

4.2 Analytic model

4.2.1 Description

Our one-period analytic model shows how the regulator's choice of emission tax depends on the cost of investment and on the fraction of firms having the opportunity to invest. With homogeneous investment costs, the regulator can choose a tax that leads to investment or a tax that does not. We show that at most only one of these will remain optimal after investment has occurred. Even with heterogeneous investment costs, the regulator faces a tradeoff between optimal investment incentives and optimal post-investment emission incentives. In selecting the tax that maximizes welfare, the regulator deviates from marginal damage pricing if that deviation produces a sufficiently better investment equilibrium.

A fraction I_0 of firms enters already having low-pollution technology. The rest of the firms must decide whether to adopt low-pollution technology (i.e., invest in low-pollution plant) or to continue with conventional technology. If they do invest, the low-pollution technology takes effect immediately, reducing the marginal benefit from emitting. The timing is that the regulator announces the tax rate, then the fraction $1 - I_0$ of firms with conventional plant decides whether to invest, and finally all firms select their emissions. Each firm is non-strategic, but the regulator can strategically induce or suppress investment. The numerical model will extend this setting to an infinite horizon while making low-pollution plant last for two periods and allowing the regulator to announce whether the tax will last for one period or for two. The following notation will also apply to the numerical model with adjustment for time subscripts.

Firms sell into distinct competitive markets, so one firm's decisions do not affect others' benefits from emitting. Additionally, firms are distributed uniformly on the unit interval, with each firm of measure 0. One firm's investment decision therefore does not by itself affect aggregate abatement cost or the regulator's policy. With conventional plant, a firm receives concave benefit $b_n(e)$ from emissions e, where $b_n(e)$ is increasing and concave: $b_n'(e) > 0, b_n''(e) < 0$. If it does not already have low-pollution technology installed, a firm can adopt it ("invest") at cost $c > 0$. Proposition 10 will extend this setting to have a continuous distribution of investment costs among firms deciding whether to invest. The benefit from emissions e with low-pollution plant is $b_i(e)$, where $b_i(e)$ is also increasing and concave. Low-pollution technology reduces the firm's benefit from additional emissions: $b_i'(e) \leq b_n'(e)\ \forall e \geq$

0.[6] Firms choose their investment probability p and emissions e to maximize their benefit from emitting net of investment cost c and net of the tax τ they pay per unit of emissions:

$$\begin{aligned} &\max_p \left\{ (1-p)\ \pi_{noinv}(\tau) + p\ \pi_{inv}(\tau) \right\} \\ =&\max_p \left\{ (1-p) \max_e \{b_n(e) - \tau e\} + p \max_e \{b_i(e) - \tau e - c\} \right\} \\ &s.t.\ e \geq 0,\ p \in [0,1] \end{aligned} \tag{4.1}$$

When firms already have low-pollution plant, they do not have the option to invest in green technology and so solve the simpler maximization problem:

$$\begin{aligned} &\pi_{lowpoll}(\tau) = \max_e \{b_i(e) - \tau e\} \\ &s.t.\ e \geq 0 \end{aligned} \tag{4.2}$$

The private benefit from adopting green technology has two components. First, because greater emissions provide less additional benefit, the firm will emit less and thereby pay less tax. Second, the net change in the "fixed" portion of profit (i.e., that portion unrelated to emission decisions) is positive in order to offset the reduced benefit form emissions. We denote this net change by Δ_0 and assume $\Delta_0 > 0$. Adopting green technology also has two costs that weigh against these benefits. First, as noted, remaining emissions provide less benefit once low-pollution technology is adopted. Second, adopting (or investing) in the low-pollution technology has cost c. The regulator does not care about firms' reduced tax burden but does also consider the net benefit of investment's effects on damages and on the optimal tax rate.

The regulator understands firms' maximization problem and seeks to maximize firms' benefit from emissions net of investment cost and net of the damage $D(e)$ caused by emis-

[6] We express matters in terms of the benefit from emitting, but this can also be understood in terms of abatement cost. We assume that low-pollution technology reduces marginal abatement cost. Note, however, that technology adoption could plausibly have other types of effects (Amir et al., 2008).

sions:

$$
\begin{aligned}
&\max_{\tau} \Bigg\{ \int_0^{I_0} \pi_{lowpoll}(\tau)\ dx + \int_{I_0}^{I_0+[1-I_0]p(\tau)} \pi_{inv}(\tau)\ dx + \int_{I_0+[1-I_0]p(\tau)}^{1} \pi_{noinv}(\tau)\ dx \\
&\qquad - D\left(\int_0^{I_0+[1-I_0]p(\tau)} e_i^*(\tau)\ dx + \int_{I_0+[1-I_0]p(\tau)}^{1} e_n^*(\tau)\ dx \right) \\
&\qquad + \tau \left[\int_0^{I_0+[1-I_0]p(\tau)} e_i^*(\tau)\ dx + \int_{I_0+[1-I_0]p(\tau)}^{1} e_n^*(\tau)\ dx \right] \Bigg\} \\
= &\max_{\tau} \left\{ w_i b_i(e_i^*) + w_n b_n(e_n^*) - p(1-I_0)c - D\left(w_i e_i^* + w_n e_n^*\right) \right\} \qquad (4.3) \\
&s.t.\ \tau \geq 0
\end{aligned}
$$

with arguments for w_i, w_n, e_i^*, e_n^*, and p suppressed in equation (4.3). The variables e_i^* and e_n^* are the emission levels chosen by firms with and without low-pollution plant, respectively. These and the investment probability p are functions of τ. The terms w_i and w_n give the fraction of firms with and without low-pollution plant once investment decisions have been made: $w_i = p(1-I_0)+I_0$ and $w_n = (1-p)(1-I_0) = 1-w_i$. Whereas the benefit functions are increasing and concave, the damage function is increasing and convex: $D'(e) > 0, D''(e) > 0$. The pollutant is a homogeneous flow pollutant, meaning damages depend neither on previous emissions nor on the distribution of emissions between firms. The convex damage function makes optimal policy sensitive to firms' adoption of low-pollution technology. With a linear damage function, marginal damage is independent of total emissions, but with a nonlinear damage function, marginal damage depends on total emissions and so on the type of plant in the fleet. The revenue from tax collection does not give the regulator a net benefit, as it merely offsets the profit lost by firms to tax payments.

4.2.2 The equilibria from which the regulator must select

We now show that there are always two candidate equilibria, where each pairs an investment level with the best tax the regulator could choose to obtain it. By virtue of moving first, the regulator must select between these two candidate equilibria. One equilibrium has firms investing and one does not, and at least one of the two equilibria includes a tax that is not ex post optimal.

Assume that the regulator sets the tax low enough that firms want to operate either type of plant. The regulator selects its tax knowing whether it will lead firms to invest. It is first of interest, however, to consider the tax the regulator would choose if it took firms' investment decisions as exogenous. This is the tax that is optimal conditional on investment. We denote it $s(I_0, p)$, recognizing that it is a function of the incoming fraction of low-pollution plant. The first proposition establishes that if the regulator does not consider the endogeneity of firms' investment, then its optimal tax decreases when more low-pollution plant ends up in the fleet. This decreased tax results from green technology reducing the marginal benefit from emitting and so reducing the total emissions generated in response to a given tax.

Proposition 1. *If the regulator takes firms' investment decisions as exogenous, then its optimal tax $s(I_0, p)$ is equal to marginal damage and decreases in both the initial fraction I_0 of low-pollution plant and the probability of investment p.*

Proof. See Appendix. □

We now know the tax the regulator would choose if it did not consider the effect on firms' incentive to invest. If firms would in fact invest with probability p for tax $s(I_0, p)$, then we have a candidate equilibrium because neither the regulator nor any firm could do better given others' decisions. The next proposition establishes that at least one candidate equilibrium in a given parameterization involves a different tax. Before presenting that result, we first show that there is only one plausible tax that makes firms indifferent between investing or not.

Lemma 2. *Assume low-pollution plant strictly reduces marginal benefit at every level of positive emissions. If there is a positive tax at which firms are indifferent between investing or not and at which firms with conventional plant would produce positive emissions, then there is only one such tax $\hat{\tau}$.*

Proof. Assume all firms operate. The incentive to invest weakly increases with the tax rate:

$$\frac{\partial\big(\pi_{inv} - \pi_{noinv}\big)}{\partial \tau} = e_n^*(\tau) - e_i^*(\tau) \geq 0 \tag{4.4}$$

using the envelope theorem. We have a strict inequality provided the conventional plant operates and provided the low-pollution plant strictly decreases marginal benefit. In this case, if there is a positive tax $\hat{\tau} > 0$ that makes $\pi_{inv} = \pi_{noinv}$, there is only one such tax. □

Proposition 3. *There are exactly two possible equilibria of tax τ and investment p decisions for any given fraction I_0 of incoming plant. At most one of these will involve a tax of $s(I_0, p)$, and if the emission benefit functions are quadratic, the other possible equilibria involve a tax in the neighborhood of the tax $\hat{\tau}$ that makes firms indifferent between investing or not.*

Proof. Proposition 1 already established that the tax $s(I_0, p)$ that the regulator would choose with exogenous investment is strictly downward-sloping in p for fixed I_0. This means that $s(I_0, 0) > s(I_0, 1)$. Lemma 2 then showed that there can be only one relevant positive tax at which firms are indifferent between investing or not. Firms invest for $s(I_0, p) > \hat{\tau}$ and do not invest for $s(I_0, p) < \hat{\tau}$. Therefore, if firms do in fact invest for $s(I_0, 1)$ as the problem with exogenous investment assumed, firms would also invest for the higher $s(I_0, 0)$. Similarly, if firms do not in fact invest for $s(I_0, 0)$ as the problem with exogenous investment assumed, firms would also not invest for the lower $s(I_0, 1)$. Finally, there is also a case which $s(I_0, 1)$ is too low to procure investment but $s(I_0, 0)$ is high enough to produce investment. In this case, neither of these tax rates is an equilibrium.

We have seen that for given I_0, at most one of $s(I_0, 0)$ and $s(I_0, 1)$ is a candidate equilibrium. We now describe the other possible equilibria. The regulator's problem with exogenous investment is globally concave if the benefit functions are quadratic. In this case, the second-order condition for the regulator's problem with exogenous investment is:

$$\begin{aligned}
&w_n\left[b_n''(e_n^*)\ e_n^{*\prime\prime}(s) + b_i'(e_i^*)\ e_i^{*\prime\prime}(s)\right] + w_i\left[b_i''(e_i^*)\ e_i^{*\prime\prime}(s) + b_i'(e_i^*)\ e_i^{*\prime\prime}(s)\right]\\
&\quad - \left[w_n e_n^{*\prime\prime}(s) + w_i e_i^{*\prime\prime}(s)\right]\ D'\big(w_n e_n^*(s) + w_i e_i^*(s)\big)\\
&\quad - \left[w_n e_n^{*\prime}(s) + w_i e_i^{*\prime}(s)\right]^2\ D''\big(w_n e_n^*(s) + w_i e_i^*(s)\big)\\
&= - \left[w_n e_n^{*\prime}(s) + w_i e_i^{*\prime}(s)\right]^2\ D''\big(w_n e_n^*(s) + w_i e_i^*(s)\big) < 0
\end{aligned} \tag{4.5}$$

where the equality uses the fact that quadratic benefits imply that $e_n^{*\prime\prime}(s) = e_i^{*\prime\prime}(s) = 0$. Because the problem is globally concave, if the regulator is constrained to only select a tax τ that is strictly less than $s(I_0, p)$, it will pick the permitted tax that is closest to $s(I_0, p)$. The same is true if it is constrained to only select a tax that is strictly higher than $s(I_0, p)$.

The constraint in our case comes from considering firms' incentive to actually invest with probability p. If $s(I_0, 0)$ is a candidate equilibrium, then it is also a candidate equilibrium for the regulator to select a tax high enough to stimulate investment (i.e., to select some $\tau > \hat{\tau}$).[7] This tax could only be an equilibrium if it is as close as possible to $s(I_0, 1)$, and because $s(I_0, 1) < \hat{\tau}$ in this example, the candidate equilibrium must have the regulator selecting $\tau = \hat{\tau} + \epsilon$ for $\epsilon > 0$ and $\epsilon << 1$. Analogously, if $s(I_0, 1)$ is a candidate equilibrium, then it is also a candidate equilibrium for the regulator to select a tax low enough to suppress investment and this tax must be $\tau = \hat{\tau} - \epsilon$. Finally, the same logic reveals that if neither $s(I_0, 0)$ nor $s(I_0, 1)$ is a candidate equilibrium, then the candidate equilibria have taxes of $\hat{\tau} + \epsilon$ and $\hat{\tau} - \epsilon$.

□

This result arises from the tension between two forces: adopting low-pollution technology drives down the optimal tax, but lower taxes provide less incentive to adopt the technology. It cannot be the case that both of the following are true: firms would invest for the lower tax set when the regulator anticipates their investment, and firms would not invest for the higher tax set when the regulator anticipates that they would not invest. In Figure 4.1a, it cannot be the case that the line $s(I_0, p)$ is below $\hat{\tau}$ for $s(I_0, 0)$ and above $\hat{\tau}$ for $s(I_0, 1)$. Instead, one of three cases must hold: 1) $\hat{\tau}$ is below $s(I_0, p)$ for all $p \in [0, 1]$, in which case $s(I_0, 1)$ is a candidate equilibrium that the regulator may choose; 2) $\hat{\tau}$ is above $s(I_0, p)$ for all $p \in [0, 1]$, in which case $s(I_0, 0)$ is a candidate equilibrium that the regulator may choose;

[7] Corollary 5 will show that the regulator would not select $\tau = \hat{\tau}$.

or 3) $\hat{\tau}$ crosses $s(I_0, p)$ for some $p \in [0, 1]$, in which case no candidate equilibrium has a tax of $s(I_0, p)$ and the regulator must either suppress or stimulate investment. The next propositions will examine the regulator's choice of equilibrium.

We can also think about this result in terms of the effect of investment on the intersection between the aggregate marginal benefit and marginal damage curves (Figure 4.1b). This figure follows the rest of the paper in assuming that when firms invest with positive probability, they rotate the aggregate marginal benefit curve downward. If marginal damage increases with pollution, this downward rotation decreases the optimal tax $s(I_0, p)$ by moving the intersection down (the optimal tax changes from point C to point A in the figure). If the break-even tax $\hat{\tau}$ is below the intersection produced when all firms invest (below point A), then firms would want to invest for any tax the regulator may choose based on these intersections. The regulator can either use that tax (point A) or suppress investment with a lower tax (point B). If $\hat{\tau}$ is above the initial aggregate marginal benefit curve's intersection with the marginal damage curve (above point C), then firms would not want to invest with any tax the regulator may choose based on these intersections. The regulator can either use the tax for that intersection (point C) or stimulate investment with a higher tax (point D). Finally, if $\hat{\tau}$ lies between these possible intersections (between points A and C), firms want to invest if the regulator assumes they will not (using the higher intersection) and do not want to invest if the regulator assumes they will (using the lower intersection). In this case, the regulator must either suppress investment (point E) or stimulate investment (point F).

4.2.3 The regulator's choice of equilibrium

We have described two candidate equilibria that ensure a tax is optimal conditional on the investment that would actually occur in response to it. Only one of these is a true equilibrium, however, because by moving first, the regulator selects the candidate equilibrium it prefers. We now consider the choice of equilibrium.

For some of the propositions, we will specialize the emission benefit and pollution damage functions to quadratic cases. The benefit from emitting with conventional plant is then $b_n(e) = b_2e^2 + b_1e + b_0$, with $b_2 \leq 0$, $b_1 \geq 0$, and $b_0 \geq 0$. The benefit from emitting with low-pollution plant becomes $b_i(e) = a_2b_2e^2 + a_1b_1e + a_0b_0$, with $a_2 \geq 1$, $a_1 \leq 1$, and $a_0 > 1$. Adoption of low-pollution technology lowers the marginal benefit from emitting, shifting it down when $a_1 < 1$ and rotating it down when $a_2 > 1$. We usually assume that $a_1 = 1$ with $a_2 > 1$, meaning that lower-pollution plant reduces marginal benefits more strongly at high emission levels. The damage caused by aggregate emissions is given by $D(e) = d_2e^2 + d_1e$ with $d_2 \geq 0$, $d_1 \geq 0$, and $b_1 \geq d_1$.[8]

[8]If a firm has conventional plant, it chooses $e_n^* = (\tau - b_1)/(2b_2)$, and if a firm has low-pollution plant, it chooses $e_i^* = (\tau - a_1b_1)/(2a_2b_2) \leq e_n^*$. The ex post optimal tax (i.e., the tax chosen with exogenous investment) is $s^* = \{a_2b_2d_1 - b_1d_2[a_2w_n + a_1w_i]\}\{a_2b_2 - d_2[a_2w_n + w_i]\}^{-1}$, where w_i and w_n are functions of I_0 and p as above. If $a_1 = 1$, the positive break-even tax is $\hat{\tau} = b_1 - 2\sqrt{-a_2b_2\Delta_0(a_2 - 1)^{-1}}$.

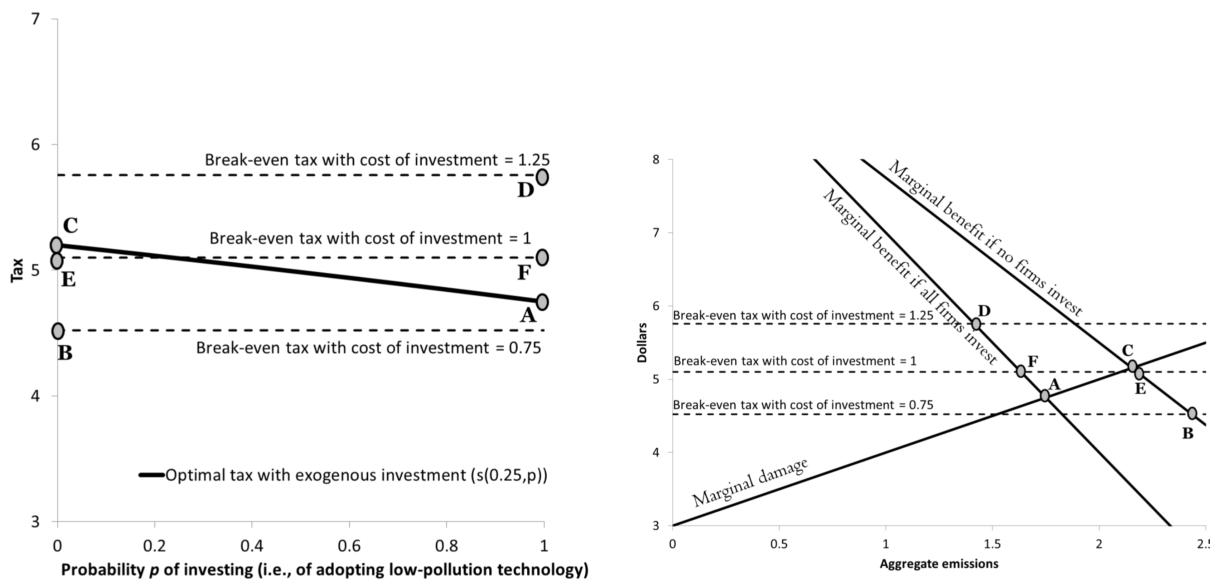
Table 4.1: Candidate equilibrium tax τ and investment probability p pairs in the one-period analytic model for a given incoming fraction I_0 of low-pollution plant. Raising investment cost c lowers Δ_0, the fixed portion of the net benefit from investment. $s(I_0, p)$ gives the tax the regulator would choose if it took the investment probability p as exogenous, and $\hat{\tau}$ is the tax that makes firms indifferent between investing or not.

Decision	Condition		Equilibria: (τ, p)	
	In terms of $\hat{\tau}$	In terms of Δ_0	No investment	Investment
Stimulate?	$s(I_0,1) < s(I_0,0) \leq \hat{\tau}$	$\Psi \equiv \frac{-a_2b_2(a_2-1)(b_1-d_1)^2}{4(a_2b_2-d_2[a_2(1-I_0)+I_0])^2} \geq \Delta_0$	$(s(I_0,0),0)$	$(\hat{\tau}+\epsilon,1)$
Suppress?	$\hat{\tau} \leq s(I_0,1) < s(I_0,0)$	$\Delta_0 \geq \frac{-a_2b_2(a_2-1)(b_1-d_1)^2}{4(a_2b_2-d_2)^2} \equiv \psi$	$(\hat{\tau}-\epsilon,0)$	$(s(I_0,1),1)$
Stimulate or suppress?	$s(I_0,1) < \hat{\tau} < s(I_0,0)$	$\psi > \Delta_0 > \Psi$	$(\hat{\tau}-\epsilon,0)$	$(\hat{\tau}+\epsilon,1)$

We can describe the conditions determining potential equilibria in terms of $\Delta_0 \equiv a_0b_0 - b_0 - c$, which is the fixed portion of the change in profit and which increases when investment cost decreases (Table 4.1 and Figure 4.2).[9] First, note that $\hat{\tau} > 0$ if and only if $\Delta_0 < -b_1^2(a_2-1)/(4a_2b_2)$. This establishes an upper bound on Δ_0, and we have also assumed $\Delta_0 > 0$ in order to make investment a potentially attractive option. When the fixed benefit from investment is low (in terms of Table 4.1, when $\Delta_0 \leq \Psi$), the possible equilibria involve either no investment or a tax high enough to stimulate investment by giving an extra push. In contrast, when the fixed benefit is high (in terms of Table 4.1, when $\Delta_0 \geq \psi$), firms will want to invest unless the regulator suppresses investment with an exceptionally low tax. Finally, for intermediate cases of the fixed benefit (when $\psi > \Delta_0 > \Psi$), the regulator's tax conditional on investment is not high enough to induce investment (because the fixed benefit is not quite high enough) but the tax rate conditional on a lack of investment is also too high to suppress investment (because the fixed benefit is not quite low enough). In this case, the regulator can only choose to stimulate or suppress investment via a tax that is suboptimally high or low once investment is realized. As either a_2 goes to 1 (so green technology has less effect) or I_0 goes to 1 (so that there are fewer opportunities for investment), then Ψ converges to ψ, removing the potential for the intermediate case.

We have identified two candidate equilibria in every parameterization, but only one of these is a true equilibrium if the regulator strictly prefers one payoff to the other. Because the regulator moves first, it can select whichever equilibrium it prefers. This ex ante regulation might not be time consistent: once investments are made, the regulator wants the tax to be $s(I_0, p)$ regardless of the tax it set ex ante. A related point will become more salient in the multiperiod numerical model when firms' adopted technology carries over into the following

[9] Requate and Unold (2003) also phrased their result in terms of the fixed cost of adopting technology.



(a) Candidate equilibria in terms of $s(I_0, p)$ (b) Candidate equilibria in terms of marginal benefit

Figure 4.1: There are always two possible equilibria that have a tax that is optimal conditional on the investment that would actually occur in response to it. When investment cost c is low (lowering the break-even tax $\hat{\tau}$), the possibilities are to obtain investment with $s(I_0, 1)$ (A) or to suppress investment with $\hat{\tau} - \epsilon$ (B). When investment cost is high, the possibilities are to forgo investment with $s(I_0, 0)$ (C) or to stimulate investment with $\hat{\tau} + \epsilon$ (B). When investment cost is intermediate, the possibilities are to suppress investment with $\hat{\tau} - \epsilon$ (E) or to stimulate investment with $\hat{\tau} + \epsilon$ (F). Both charts use $I_0 = 0.25$ and are parameterized as in the numeric model (see Table 4.2) but with $a_2 = 1.5$.

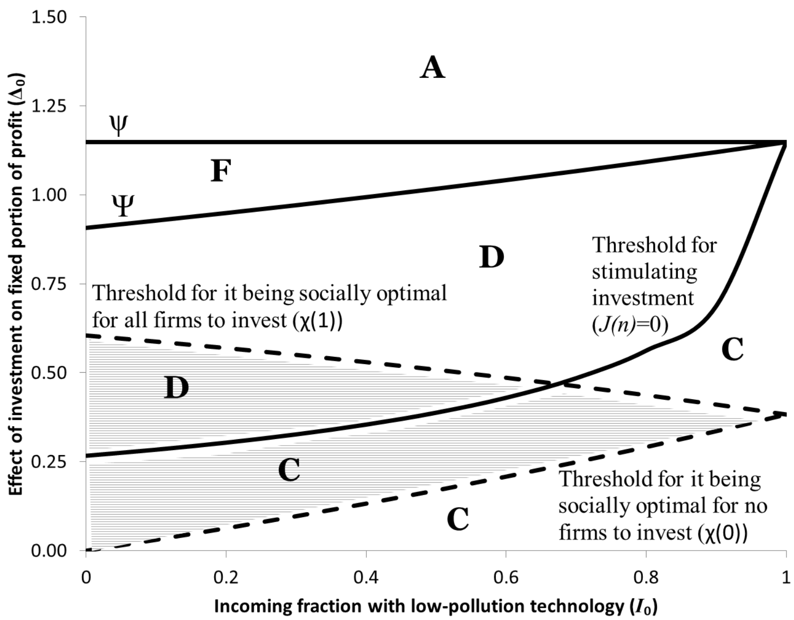


Figure 4.2: The regulator's chosen equilibrium depends on I_0 and Δ_0, where Δ_0 increases as investment cost c falls. The letters and calibration correspond to those in Figure 4.1, and ψ and Ψ are as in Table 4.1. When investment cost c is low ($\Delta_0 \geq \psi$), the regulator always obtains investment with $s(I_0, 1)$ (A). When investment cost is intermediate ($\Psi < \Delta_0 < \psi$), the regulator stimulates investment with $\hat{\tau} + \epsilon$ (F). When investment cost is high ($\Delta_0 \leq \Psi$), the regulator either stimulates investment with $\hat{\tau} + \epsilon$ (D) or forgoes investment with $s(I_0, 0)$ (C). The shaded area shows where it is socially optimal for firms to invest with probability $p \in (0, 1)$. Above the shaded area, it is socially optimal for all firms to invest, and below the shaded area, it is socially optimal for no firms to invest. $J(n)$ and $\chi(p)$ are defined by equations (4.24) and (4.29) in the appendix.

period and the regulator can pick a tax based on the new incoming fraction of plants having low-pollution technology.

In considering which equilibrium the regulator chooses to play, we begin with the case in which Δ_0 is strictly between Ψ and ψ. We have seen that the regulator will set the tax within ϵ of firms' break-even tax $\hat{\tau}$, with a tax just above $\hat{\tau}$ acting to stimulate investment and a tax just below $\hat{\tau}$ acting to suppress investment. The regulator in this case would never choose to suppress investment:

Proposition 4. *When $\Psi < \Delta_0 < \psi$, the regulator always chooses to stimulate investment with a tax of $\hat{\tau} + \epsilon$.*

Proof. We saw in Proposition 3 that the regulator in this situation will choose a tax within ϵ of $\hat{\tau}$. The regulator compares the welfare $W_{w_i=1}$ from the equilibrium with investment to welfare $W_{w_i=I_0}$ from the equilibrium without investment:

$$
\begin{aligned}
W_{w_i=1} =& (1-I_0)\pi_{inv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D(e_i^*(\hat{\tau})) + \hat{\tau}\ e_i^*(\hat{\tau}) \qquad (4.6)\\
W_{w_i=I_0} =& (1-I_0)\pi_{noinv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D\left((1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right)\\
&+ \hat{\tau}\left[(1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right]\\
=& (1-I_0)\pi_{inv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D\left((1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right)\\
&+ \hat{\tau}\left[(1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right] \qquad (4.7)
\end{aligned}
$$

using the fact that $\hat{\tau}$ is defined as the tax that makes firms indifferent between investing or not. Assume the regulator finds it optimal to suppress investment by choosing $\hat{\tau} - \epsilon$:

$$
\begin{aligned}
& W_{w_i=I_0} > W_{w_i=1}\\
\Rightarrow\ & D(e_i^*(\hat{\tau})) > D((1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau}))\\
\Rightarrow\ & e_n^*(\hat{\tau}) < e_i^*(\hat{\tau}) \qquad (4.8)
\end{aligned}
$$

However, we saw in proving Proposition 1 that $e_n^*(\hat{\tau}) \geq e_i^*(\hat{\tau})$. This contradicts the condition in equation (4.8) for suppressing investment. The regulator therefore finds it optimal to stimulate investment with $\hat{\tau} + \epsilon$.

□

When the regulator must either stimulate or suppress investment with a tax within ϵ of the critical level, it always chooses to stimulate investment. Being in the situation with $s(I_0, 1) < \hat{\tau} < s(I_0, 0)$ means that $\hat{\tau}$ is the regulator's preferred tax among those that could actually obtain each possible investment outcome, and because $\hat{\tau}$ is defined such that it equalizes firms' profits between investment scenarios, the only difference in welfare between using $\hat{\tau} + \epsilon$ and $\hat{\tau} - \epsilon$ is in damages. When there is no change in firm profits, the regulator chooses the tax that causes fewer damages, which is the tax that causes fewer emissions by inducing investment into low-pollution plant. Corollary 5 describes the implications of this same logic for the plausibility of firms employing mixed strategies in equilibrium:

Corollary 5. *The regulator never wants to induce firms to invest with probability $p \in (0, 1)$.*

Proof. In order to induce investment with probability $p \in (0, 1)$, the regulator would have to set the tax at $\hat{\tau}$ because firms are then indifferent between investing or not. While any p is in fact compatible with such a tax, suppose the regulator could choose which p firms select when they see $\hat{\tau}$.

First, assume that $I_0 = 1$. Because no firms have an investment decision to make, the regulator has one optimal tax that holds irrespective of whether it believes firms would want to invest (i.e., $s(1, 0) = s(1, 1)$). It would then only select a tax within ϵ of $\hat{\tau}$ if the optimal tax happens to lie in that region, not in order to stimulate or suppress investment as in other cases.

Next assume $I_0 < 1$. Consider any two values p_1 and p_2 such that both are probabilities and $p_1 < p_2$. When the tax is at $\hat{\tau}$, firms' profits are the same regardless of whether they invest or not, so the regulator decides between p_1 and p_2 solely on the basis of the pollution damages they cause. Because p_2 results in a greater fraction of low-pollution plant, it also results in lower emissions and reduced damage. The regulator therefore chooses p_2. Because this applied to arbitrary p_1 and p_2 on the unit interval, the regulator therefore always chooses $p = 1$ when $I_0 < 1$ and the tax is $\hat{\tau}$. Furthermore, it can ensure this investment outcome by setting the tax at $\hat{\tau} + \epsilon$.

□

We have seen that in equilibrium either all firms invest or none invest, and the regulator prefers stimulating investment to suppressing investment with a tax in the neighborhood of $\hat{\tau}$. Next, consider the case with Δ_0 high (i.e., investment cost is low) in which the regulator either selects $s(I_0, 1)$ or suppresses investment with $\hat{\tau} - \epsilon$. Because neither $s(I_0, 1)$ nor $\hat{\tau}$ depend on I_0, these two equilibria are possible for all I_0 or for none. Does the regulator ever choose to suppress investment, and if so, under what conditions? In fact, the regulator would never choose to suppress investment with the lower tax of $\hat{\tau} - \epsilon$:

Proposition 6. *When $\psi \leq \Delta_0$, the regulator never chooses to suppress investment. Investment therefore always occurs, with the regulator selecting the tax $\tau = s(I_0, 1)$.*

Proof. Define $s \equiv s(I_0, 1)$ and assume $s > \hat{\tau}$. The regulator compares the welfare $W_{w_i=1}$ from the equilibrium with investment with welfare $W_{w_i=I_0}$ from the equilibrium without investment:

$$\begin{aligned} W_{w_i=1} =& \pi_{lowpoll}(s) - (1-I_0)c - D(e_i^*(s)) + s\ e_i^*(s) && (4.9)\\ W_{w_i=I_0} =& (1-I_0)\pi_{noinv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D((1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau}))\\ & + \hat{\tau}\left[(1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right]\\ =& (1-I_0)\pi_{inv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D((1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau}))\\ & + \hat{\tau}\left[(1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right]\\ =& \pi_{lowpoll}(\hat{\tau}) - (1-I_0)c - D((1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau}))\\ & + \hat{\tau}\left[(1-I_0)e_n^*(\hat{\tau}) + I_0 e_i^*(\hat{\tau})\right] && (4.10) \end{aligned}$$

using the fact that $\hat{\tau}$ is defined as the tax that makes firms indifferent between investing or not. The regulator chooses to suppress investment with $\hat{\tau} - \epsilon$ if and only if:

$$W_{w_i=I_0} > W_{w_i=1} \tag{4.11}$$

Assume the regulator suppresses investment, so $W_{w_i=I_0} > W_{w_i=1}$. Let $W_{\hat{\tau}+\epsilon}$ and $W_{\hat{\tau}-\epsilon}$ be the regulator's value from stimulating investment with a tax of $\hat{\tau}+\epsilon$ and suppressing investment with a tax of $\hat{\tau}-\epsilon$ (as in the scenario with $\Psi < \Delta_0 < \psi$). We know from Proposition 4 that $W_{\hat{\tau}+\epsilon} > W_{w_i=I_0} = W_{\hat{\tau}-\epsilon}$. Because $W_{w_i=1}$ uses the tax $s(I_0, 1)$ that is optimal if investment were exogenous, we also know that $W_{w_i=1} \geq W_{\hat{\tau}+\epsilon}$. In all, this gives $W_{w_i=I_0} > W_{w_i=I_0}$, a contradiction. Therefore the regulator never suppresses investment.

The appendix has an alternate proof using quadratic benefit and damage functions.

□

The regulator never chooses to suppress investment, but before we discuss the intuition, the next proposition shows that the regulator may in fact stimulate investment if Δ_0 is not too far below Ψ. Moreover, for Δ_0 near Ψ, the regulator stimulates investment for low I_0 if at all.

Proposition 7. *When $\Delta_0 \leq \Psi$ and benefits and damages are quadratic, there is a range of Δ_0 bounded above by Ψ in which the regulator would choose to stimulate investment with a tax of $\hat{\tau}+\epsilon$ for some I_0. Further, for Δ_0 in the top portion of this range, the I_0 for which the regulator stimulates investment form a connected set bounded below by 0.*

Proof. See appendix. □

We have shown that the regulator will sometimes stimulate investment but will never suppress it. The reason is that to do either involves setting a tax within ϵ of the break-even tax $\hat{\tau}$. In this region, the regulator would always rather have investment than not because firms' profits are equalized and the only welfare difference is in pollution damages. These damages are lower with investment in lower-pollution plant. It would never be optimal to suppress investment because the regulator could do better with a tax in the same neighborhood. Now consider the option of stimulating investment. If the regulator compares the best tax in this neighborhood of $\hat{\tau}$ to a tax $s(I_0, 0)$ that would not obtain investment, it faces a trade-off between using a tax that is optimal conditional on not having investment (i.e., a tax of $s(I_0, 0)$) and using a tax that obtains lower-pollution plant but is less adapted to its investment outcome (i.e., a tax of $\hat{\tau} + \epsilon$). If $s(I_0, 0)$ is near to $\hat{\tau}$, then stimulating looks more attractive because it involves a relatively small change in tax for the gain in emissions. However, if $s(I_0, 0)$ is well below the break-even tax, then increasing the tax to $\hat{\tau} + \epsilon$ carries a greater cost.

4.2.4 The social optimality of equilibrium outcomes

We already found the socially optimal tax $s(I_0, p)$ conditional on a probability p of investing. We now consider the socially optimal level of investment. The next proposition shows that it is often socially optimal for investment to occur with probability $p \in (0, 1)$, even though Corollary 5 showed that this never happens in equilibrium.

Proposition 8. *Assume benefits and damages are quadratic. It is socially optimal for firms to invest with probability $p \in (0, 1)$ for some connected set of Δ_0, and when $\Delta_0 = \Psi$, it is socially optimal for firms to invest with positive probability. Further, if $\Delta_0 = \Psi$ and either $a_2 \leq 2$ or I_0 is sufficiently large, then it is socially optimal for all firms to invest with probability 1.*

Proof. See Appendix. □

Corollary 9. *If $d_2 \geq -b_2/2$ and benefits and damages are quadratic, it is always socially optimal for there to be some low-pollution plant in the fleet.*

Proof. See Appendix. □

In Figure 4.2, the only regions in which the equilibrium outcome is socially optimal are region A and the part of region C that is below the dashed line corresponding to $\chi(0)$. For the part of region C in the shaded area, the socially optimal outcome has investment with some positive probability $p \in (0, 1)$, but the equilibrium outcome has no investment. The part of region C that lies above the shaded area obtains no equilibrium investment when it would have been socially optimal for all firms to invest. In region D, the regulator chooses to stimulate investment with a tax of $\hat{\tau} + \epsilon$. However, no part of region D achieves the socially optimal emission outcome, and the part of region D in the shaded area produces too much investment since the social optimum has some firms not investing. Both region F and the part of region D above the line $\chi(1)$ achieve the socially optimal investment outcome but do so with a tax that is higher than optimal. Finally, all of region A achieves the socially optimal outcome of full investment with a tax of $s(I_0, 1)$. The outcomes therefore are socially optimal for region A and the lowest part of region C, have inefficiently low investment in the two upper parts of region C, have both inefficiently high investment and an inefficiently high tax in the lower part of region D, and have an inefficiently high tax in region F and the upper part of region D.

Two aspects of the ex ante tax instrument tend to produce inefficient outcomes. First, this instrument cannot obtain investment from only a fraction of firms because all firms always have the same incentive to invest. The linear tax does not respond to realized investment and firms are homogeneous, leading investment outcomes to be all-or-nothing except for a knife-edge case at which anything is possible and which the regulator always finds suboptimal. The shaded area in Figure 4.2 indicates the region in which it would be socially efficient for only a fraction of firms to invest, but the regulator cannot obtain this outcome. If the regulator used the tax that would be optimal with only a fraction $p \in (0, 1)$ of firms investing, then either all firms or no firms would invest. When constrained by outcomes that might plausibly happen, the regulator's ex post optimal tax must therefore be $s(I_0, 0)$ or $s(I_0, 1)$, but these taxes often do not supply the investment incentives to support the investment outcomes they assume.

However, even if firms were heterogeneous so that the regulator could obtain investment from only some of them, there is a second aspect of the ex ante linear tax that tends to produce socially inefficient outcomes: the regulator chooses the tax while cognizant that it is fixing both investment incentives and emission incentives, but the single instrument cannot generally accomplish both goals efficiently. We can see this by making investment costs heterogeneous.[10] Because the regulator can now choose exactly how many firms invest, this extension removes the first aspect of inefficiency discussed above. However, the ex ante linear tax still often produces socially inefficient outcomes:

Proposition 10. *Assume investment costs have a continuous distribution among firms with the option to invest and that this distribution has full support among positive real numbers. As long as $I_0 < 1$ and the break-even tax does not increase one-for-one with investment cost, the regulator's chosen ex ante tax will not obtain socially optimal investment along with socially optimal emissions conditional on investment.*

Proof. See Appendix. □

If the regulator could set one tax prior to firms investing and a second after they invested, and if firms did not anticipate this second tax, then the regulator could use the first to achieve optimal investment and the second to obtain optimal emissions. These two taxes will not generally be the same, so the ex ante tax faces a tradeoff between them. In a setting with heterogeneous investment costs, the regulator weights the tax between these other two taxes and sacrifices some efficiency in both investment and emissions conditional on investment. In a setting with identical investment costs, the regulator cannot smoothly

[10] Investment costs could be heterogeneous because firms or low-pollution technologies are heterogeneous. This model is equivalent to one in which the low-pollution technologies have a heterogeneous effect on the fixed portion of profit but an identical effect on marginal abatement cost.

adjust investment outcomes and so the regulator more often faces a choice between optimal investment and optimal emissions conditional on investment. The ability of the regulator to stimulate investment with a higher tax increases the region over which realized investment is optimal relative to a world in which the regulator were restricted to using the ex post optimal tax $s(I_0, p)$, but it does so by using a tax that is inefficiently high. In Figure 4.1b, this tax is above the intersection of realized aggregate marginal benefit and marginal damage. In some cases, however, the loss from the tax inefficiency can be less than the loss from obtaining the inefficient investment outcome with a tax that is conditionally efficient. Firms often tend to under-invest because they value investment at the current tax rate but do not value its effect in changing the tax rate. In equilibrium, firms must find investment optimal at the tax the regulator would choose in anticipation of them investing, but this tax is lower (and so provides fewer incentives to invest) than the tax the regulator would use if they did not invest. Being able to lower the tax provides its own gains in terms of profits, but this change does not enter firms' calculations. Because firms move after the regulator has committed to the tax rate, they take the tax rate as given when deciding upon their investments. The regulator can only compensate for this inefficiency by using a tax that introduces its own inefficiency. The regulator could consistently achieve efficiency if it used a second instrument to control investment or used a policy like tradable permits that conditions the incentive to invest on the investment that occurs.

4.2.5 Comparison to a quantity instrument

Previous work has shown that a quantity instrument can obtain the socially optimal level of investment in cases where a price instrument might not (Requate and Unold, 2003; Karp, 2008). Two relevant differences between a quantity policy and a price policy cause this result. First, with a given tax, all firms have the same incentive to invest, and that incentive does not depend on how many other firms invest.[11] Firms' strategies are therefore independent of each other, and each invests with probability 1 or 0 except for a knife-edge tax at which it will invest with any probability. Further, with identical firms, aggregate investment will be an all-or-nothing outcome. In contrast, for a given quantity policy with tradable permits (i.e., for a given cap on aggregate emissions), the cost of emitting (and so the incentive to invest) depends on other firms' investment decisions because these affect the market price of pollution permits. Even when the cap is set prior to investment, the cost of emissions is automatically conditioned on investment decisions in a way unavailable with the simplest linear tax (Requate and Unold, 2003).[12] Firms' strategies become strategic substitutes because, by lowering the permit price, investment by one firm reduces other firms' incentive to invest (Karp, 2008). The investment outcome with a quantity policy and identical firms need

[11] Note that this claim may not hold if, for instance, firms sell into the same noncompetitive product markets.

[12] A more complex tax could perform better. For instance, a tax could be a nonlinear function of aggregate emissions or could be conditioned on observed or reported investments (e.g., Dasgupta et al., 1980).

not be all-or-nothing; because firms can play mixed strategies in equilibrium, it is possible that only a fraction of the firms invest in response to a quantity policy.

There is a second relevant difference between a quantity policy and a price policy, involving how the chosen level of the policy depends on anticipated investment. With a price policy, if the regulator anticipates adoption of low-pollution technology, it wants to choose a lower tax as long as marginal damage increases in emissions. With a quantity policy, it wants to move to the same new intersection of marginal damage and marginal abatement cost if it anticipates adoption of low-pollution technology, but it does so by lowering the cap on emissions. For a given level of investment, this raises the cost of emitting, whereas lowering the tax reduces the cost of emitting. The permit price is eventually reduced to the level of the tax, but only because firms invest in low-pollution technology; with a tax, the cost of emissions is set at the lower level whether or not firms adopt the low-pollution technology. The policies have the same emission outcome if investment is exogenous, but if investment is endogenous and occurs after the policy is set, the lowered tax reduces the incentive to invest while the lowered cap increases the initial incentive to invest but investment then makes the permit price fall.

In sum, with a price policy, firms' incentives to invest do not depend on other firms' investments and the regulator reduces the incentive to invest when it anticipates investment. With a quantity policy, firms' incentives to invest decrease as other firms invest and the regulator increases the incentive to invest when it anticipates investment.

4.3 Multiperiod numeric model

4.3.1 Description and solution procedure

We have seen how endogenous investment leads the regulator to adjust its tax rate in a one-period setting. If damages are convex, it is not the case that the modeled regulator could pick a tax without considering firms' investment response and trust that the optimal level of investment will follow. We now extend these results to an infinite horizon setting using quadratic benefit and damage functions. As before, the low-pollution plant is available for use in the same period in which the investment is made, but now low-pollution technology lasts two periods. Also, the regulator and firms maximize welfare and profit over all time rather than over only a single period. Finally, the regulator can now choose whether a tax will last for one period or for two, where a two-period tax means that it will have no choice to make in the subsequent period.

This new setting is best represented as a dynamic programming problem. Let I_{t-1} indicate the fraction of firms in period t who already have low-pollution plant and so have no investment decision to make. Also, let γ^i_{t-1} be an indicator variable equal to 1 when firm i invested in low-pollution plant in period $t-1$, which means firm i already has low-pollution plant in period t. Let z_t be the duration of the regulator's tax in period t. A just-announced two-period tax has $z_t = 2$, while both a just-announced one-period tax and a carried-over two-period tax have $z_t = 1$. The timing within a period is that the regulator selects the time t tax rate and its duration, followed by firms simultaneously selecting their time t investments and then simultaneously selecting their time t emissions. We model a flow pollutant, so firms' emission decisions in period t do not affect pollution damages in future periods. The regulator and firms all have rational expectations with regards to their own actions and to others' actions.

Because they move after the regulator in time t, firms take the time t policy as given. They make their investment decisions partly with an eye on future policies, and their optimal emissions depend on the type of plant they have (as determined by investment decisions in the current period and in the previous one). Firm i's problem is to maximize the present value of the stream of profits conditional on the regulator's current and future tax rates:

$$
\begin{aligned}
&V^i(\tau_t, z_t, I_{t-1}, \gamma^i_{t-1}) = \\
&\quad \begin{cases} \max_{p^i_t} \{(1-p^i_t)\ \pi_{noinv}(\tau_t) + p^i_t\ \pi_{inv}(\tau_t)\} + \beta_f\ E[V^i(\tau_{t+1}, z_{t+1}, I_t, \gamma^i_t)] & \text{if } \gamma^i_{t-1} = 0 \\ \pi_{lowpoll}(\tau_t) + \beta_f\ E[V^i(\tau_{t+1}, z_{t+1}, I_t, 0)] & \text{if } \gamma^i_{t-1} = 1 \end{cases} \\
&s.t.\ e^i_t \geq 0,\ p^i_t \in [0,1]
\end{aligned} \tag{4.12}
$$

where i superscripts specialize variables to firm i, π terms give the firm's profit after optimally selecting emissions and are defined as in the analytic model, and $0 < \beta_f < 1$ gives a firm's per-period discount factor. The V^i terms on the right-hand side of the equation give the continuation value, which depends on the next period's tax rate and duration as chosen by the regulator, on the total fraction of low-pollution plant entering the next period as determined by firms' investment decisions (i.e., $I_t = p(1 - I_{t-1})$), and on whether firm i will have low-pollution plant entering the next period ($\gamma^i_t = 1$ if $p^i_t = 1$ with $\gamma^i_{t-1} = 0$, and $\gamma^i_t = 0$ if either $p^i_t = 0$ or $\gamma^i_{t-1} = 1$). The firm's continuation value depends on other firms' investment decisions via I_t, and the firm assumes other firms will invest as it does. In equilibrium, the firm chooses to invest with probability p_t given that other firms invest with probability p_t. If the firm enters the period with low-pollution plant ($\gamma^i_{t-1} = 1$), its current actions cannot affect its continuation value as its only optimization decision is over time t emissions. Because this is an infinite-horizon model, the value function is the same in each period. The expectation operator is over firms' investments, which affect both γ^i_t and I_t, and accounts for firms playing mixed strategies.

The regulator's problem is to maximize the present value of the benefit from emissions net of investment cost and pollution damage. The regulator knows how firms will react to its announced tax and duration and chooses both with an eye to their effect on both investment and pollution. Its value function W has arguments indicating the regulator's choices in the previous period about today's tax rate (z_{t-1} and τ_{t-1}) and indicating the incoming fraction of low-pollution plant I_{t-1}:

$$W(\tau_{t-1}, z_{t-1}, I_{t-1}) =$$
$$\begin{cases} \max_{\tau_t, z_t} \Big\{ \int_0^{I_{t-1}} \pi_{lowpoll}(\tau_t)\ dx + \int_{I_{t-1}}^{I_{t-1}+[1-I_{t-1}]p_t} \pi_{inv}(\tau_t)\ dx + \int_{[1-I_{t-1}]p_t}^{1} \pi_{noinv}(\tau_t)\ dx \\ \quad -D\left(\int_0^{I_{t-1}+[1-I_{t-1}]p_t} e_i^*(\tau_t)\ dx + \int_{I_{t-1}+[1-I_{t-1}]p_t}^{1} e_n^*(\tau_t)\ dx \right) & \text{if } z_{t-1} = 1 \\ \quad +\tau_t \left[\int_0^{I_{t-1}+[1-I_{t-1}]p_t} e_i^*(\tau_t)\ dx + \int_{I_{t-1}+[1-I_{t-1}]p_t}^{1} e_n^*(\tau_t)\ dx \right] \\ \quad +\beta_r\ E[W(\tau_t, z_t, I_t)] \Big\} \\ \int_0^{I_{t-1}} \pi_{lowpoll}(\tau_{t-1})\ dx + \int_{I_{t-1}}^{I_{t-1}+[1-I_{t-1}]p_t} \pi_{inv}(\tau_{t-1})\ dx + \int_{[1-I_{t-1}]p_t}^{1} \pi_{noinv}(\tau_{t-1})\ dx \\ \quad -D\left(\int_0^{I_{t-1}+[1-I_{t-1}]p_t} e_i^*(\tau_{t-1})\ dx + \int_{I_{t-1}+[1-I_{t-1}]p_t}^{1} e_n^*(\tau_{t-1})\ dx \right) & \text{if } z_{t-1} = 2 \\ \quad +\tau_{t-1} \left[\int_0^{I_{t-1}+[1-I_{t-1}]p_t} e_i^*(\tau_{t-1})\ dx + \int_{I_{t-1}+[1-I_{t-1}]p_t}^{1} e_n^*(\tau_{t-1})\ dx \right] \\ \quad +\beta_r\ E[W(\tau_{t-1}, 1, I_t)] \end{cases}$$
$$s.t.\ \tau_t \geq 0,\ \ z_t \in \{0, 1\}$$

where p_t represents the investment probability chosen by the fraction $(1 - I_{t-1})$ of identical firms making an investment decision. The regulator's per-period discount factor is β_r. The regulator at time t chooses its tax policy if and only if $z_{t-1} = 1$; otherwise, it already set its time t tax policy at time $t-1$ (giving $\tau_t = \tau_{t-1}$ and $z_t = 1$) and has no maximization problem because it has no choice to make. As above, the expectation operator is over aggregate firm investments and is relevant when firms play mixed strategies. Firms' profits are functions of the current period's tax rate, and firms' investment probabilities are affected by the current tax rate and its duration.

We calibrate the numeric model as in Table 4.2, running it with several investment costs in order to compare investment outcomes and the regulator's decisions over a range of settings. We solve the model by starting in a terminal period and stepping backward until the policy rules converge.[13] Once the policy rules cease to change from period to period, we have moved beyond the influence of the terminal period. We solve for the optimal actions and value at a grid of nodes in the state space, and we approximate the value functions and the tax policy rule over the remainder of the state space using linear spline basis functions. When evaluating the duration policy, we use nearest-neighbor interpolation. We do not interpolate equilibrium investment policy but instead always calculate it exactly using the approximated value functions and other policy rules.

Within period t, we begin by solving for the value function of a firm with the option to invest. For a given time t tax and duration, this value function only depends on future tax and investment choices, which we have already solved by stepping backward through time,

[13]We define convergence by policy rules rather than value functions because, in the backward iteration, earlier value functions will be affected by including more periods in their continuation value.

Table 4.2: Parameterization of the numeric model

Parameter	Value	Description
b_2	-2	Quadratic term in emission benefit function
b_1	10	Linear term in emission benefit function
b_0	1	Constant term in emission benefit function
a_2	1.75 [a]	Multiplier for b_2 due to low-pollution technology
a_1	1	Multiplier for b_1 due to low-pollution technology
a_0	3	Multiplier for b_0 due to low-pollution technology
c	1.22,1.28	Cost of adopting low-pollution technology (i.e., investment cost)
d_2	1	Quadratic term in damage function
d_1	3	Linear term in damage function
d_0	0	Constant term in damage function
β_f	0.95	Firms' discount factor
β_r	0.95	The regulator's discount factor

[a] In the figures that illustrate the analytic results, we use $a_2 = 1.5$. Because low-pollution plant only lasts for one period in the analytic model, the emission benefits after investment need to be greater in order to offset a given investment cost.

and on the time t investment choices of other firms, which are all identical and are determined by the need for investment to be an equilibrium outcome. Investing with probability $\tilde{p}$ is an equilibrium outcome if and only if a firm cannot do better by investing with any probability $p \neq \tilde{p}$ given that other firms invest with probability $\tilde{p}$. We must solve for these firms' value functions first because the continuation values for all actors depend on the fraction of firms that adopt low-pollution technology in the current period. We next solve for the value function of a firm that invested in the period immediately before and so already has low-pollution plant. This firm's value function depends on the chosen tax and duration for the current period and on future actions by it and others. We can calculate its continuation value since we already know all value functions for the following period and since we know the firm's state variables in the next period by using the regulator's policy rules in the next period and the current period's equilibrium investment. Finally, we solve for the regulator's value function when it can and cannot choose the current period's tax. This solution is enabled by already knowing how firms will respond in this period to each possible choice of tax and duration. Once we have each actor's optimal actions and value function in a given period, we repeat the procedure in the preceding period with the new continuation values and next-period actions.

4.3.2 Numeric results

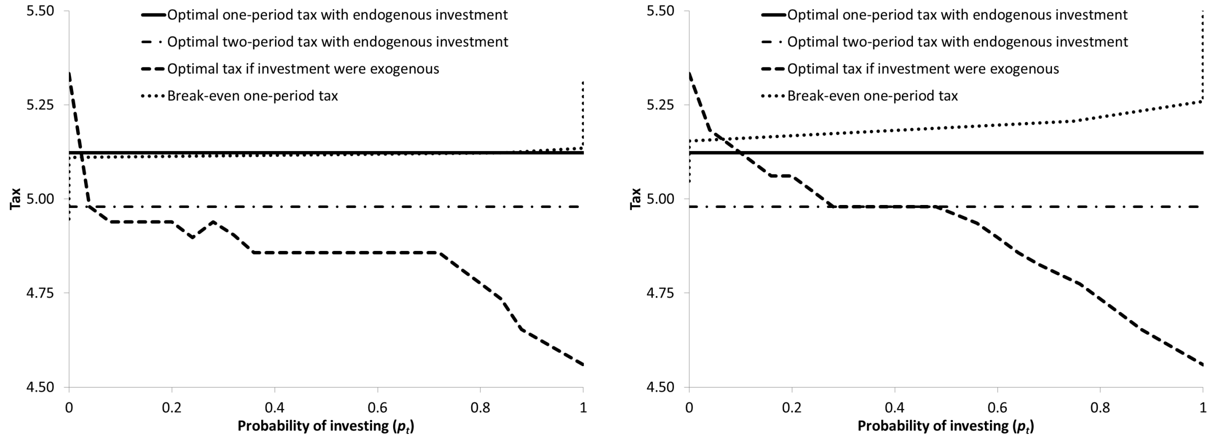
The numeric model demonstrates three results. First, for an intermediate range of investment cost, only a fraction of firms invest in equilibrium. In contrast to the one-period setting, this occurs even when firms have identical investment costs. Second, the regulator only chooses to lock in the tax rate for two periods for a narrow range of investment cost and only when there is little low-pollution plant already in the fleet. Third, the emission outcomes are often socially suboptimal because the regulator still lacks the ability to separately incentivize investment and emissions.

The first result concerns the regulator's ability to obtain investment from only a fraction of firms. In the one-period setting with homogeneous investment costs, the ex post optimal tax slopes down in the investment probability p but the regulator can only select a tax that is constant in p (Figure 4.1a). To be optimal, this tax must be at one of the end points of the downward-sloping ex post optimal tax (i.e., at $s(I_0, 0)$ or at $s(I_0, 1)$) or it must be within a small neighborhood of the "break-even" tax $\hat{\tau}$ that makes firms indifferent between investing or not. In the multiperiod setting, the regulator must still select a tax that holds for all investment probabilities (i.e., the tax must still be constant in p_t for a given I_{t-1}), but because low-pollution plant lasts for two periods, firms' investment incentives also depend on the time $t+1$ choice of tax. For a given time t incoming plant fraction I_{t-1}, the time $t+1$ incoming plant fraction I_t is increasing in the time t investment probability p_t: $I_t = p_t\ (1 - I_{t-1})$. The time $t+1$ tax should therefore decrease in p_t. Because they expect lower p_t to lead to a higher τ_{t+1}, a firms is more willing to invest if it believes other firms will not invest. The regulator's future ability to partially respond to this period's investment allows the investment incentive to be imperfectly conditioned on realized investment.

In the one-period setting, the break-even tax $\hat{\tau}$ that makes firms indifferent between investing or not is constant in the probability with which firms invest. Other firms' investments have no chance to influence firm i's investment decision. However, in the multiperiod setting, the break-even one-period tax $\hat{\gamma}_t$ does depend on the probability with which other firms invest because it also depends on the tax that firms expect the regulator to select at time $t+1$. If many other firms invest at time t (i.e., if p_t is high), then firms expect a relatively low time $t+1$ tax and so will need a relatively higher time t tax if they are to profit from investment (i.e., $\hat{\gamma}_t$ must be relatively high for p_t high). Conversely, if few other firms invest at time t (i.e., if p_t is low), then firms expect a relatively high time $t+1$ tax and so will invest for a relatively low time t tax because they can recover more of the investment cost in the following period (i.e., $\hat{\gamma}_t$ is relatively low for p_t low). The break-even one-period tax $\hat{\gamma}_t$ can therefore slope upward over some range of p_t in the multiperiod setting.

In both the multiperiod and the one-period setting, the chosen time t tax τ_t must be constant in the investment probability p_t even as the ex post optimal tax s is downward-sloping in p_t. When s is completely above the break-even tax, the ex post optimal tax with full investment is a candidate equilibrium, and when s is completely below the break-even tax, the ex post optimal tax with no investment is a candidate equilibrium. There is a crucial difference, however, due to the differing slopes of the break-even taxes between these settings. In the one-period setting, the regulator can potentially completely suppress or stimulate investment with a small change in the ex post optimal tax because $\hat{\tau}$ is constant in p_t, but in the multiperiod setting, the positive slope of $\hat{\gamma}_t$ in p_t means that a small change in the tax only incrementally shifts the fraction of firms that invest. The regulator is therefore less likely to abandon the ex post optimal tax for a tax that obtains an extreme investment outcome. Instead, as in the setting with heterogeneous investment costs, the regulator is more likely to abandon the ex post optimal tax for a tax that causes a smaller change in investment. Finally, when s crosses $\hat{\tau}$ in the one-period setting, the regulator selects a tax within ϵ of $\hat{\tau}$ to obtain full or no investment. In the multiperiod setting, when s crosses $\hat{\gamma}_t$, the ex post optimal tax is consistent with a particular fraction of firms investing, not with any fraction of firms investing. A given τ_t is now associated with a single mixed strategy outcome because equilibrium investment probabilities are constrained by the regulator's response in the next period. Because the fraction of firms investing changes continuously with the time t tax rate, the regulator cannot follow the logic of Proposition 4 in obtaining any desired level of investment with the same tax rate. The regulator may now consider a tax further than ϵ from the break-even level even though it obtains less additional investment than did the change of ϵ in the one-period setting. If the regulator chooses a tax other than the ex post optimal tax that crosses $\hat{\gamma}_t$, it changes the tax by a larger magnitude than in the one-period setting but obtains a smaller investment benefit.

We now look at the results of the model run with investment costs of 1.22 and 1.28. In doing so, we learn about the situations in which the regulator chooses a two-period policy. We choose these investment costs because they produce interesting outcomes involving two-

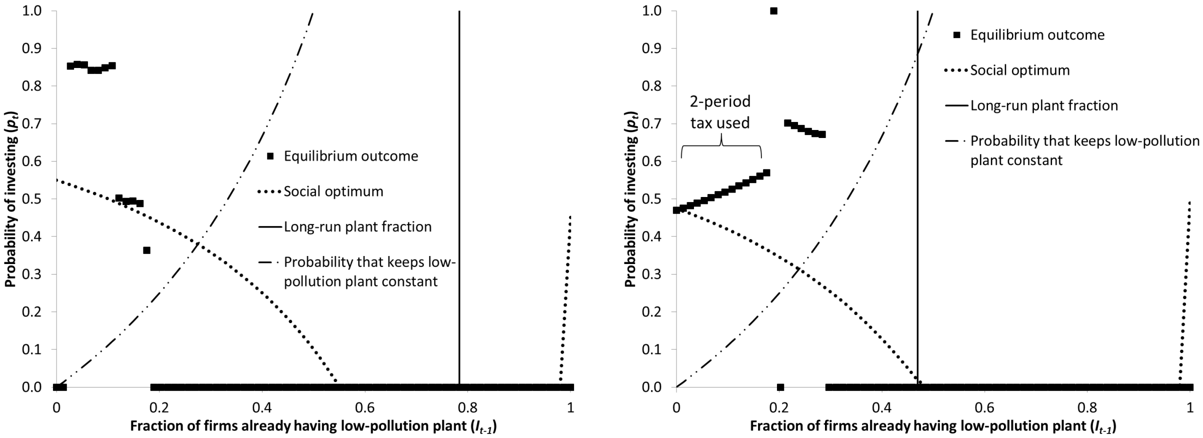


(a) Investment cost: 1.22 (b) Investment cost: 1.28

Figure 4.3: The relation between the probability p_t that firms invest and the optimal one-period and two-period taxes when there is no low-pollution plant already in the fleet ($I_{t-1} = 0$). Also shows the one-period tax $\hat{\gamma}$ that makes firms indifferent between investing or not and the one-period tax s that would be optimal if investment occurred exogenously with probability p_t.

period policies and mixed strategies. With the lower investment cost, the regulator always uses a one-period policy, and the post-investment fraction of low-pollution plant quickly jumps to 0.78 and stays there forever. With the higher investment cost, the regulator selects a two-period policy when less than 20% of firms already have low-pollution plant, and the post-investment fraction of low-pollution plant quickly jumps to 0.47 and stays there forever. Note that all of the following results are merely indicative as we are still obtaining convergence in our solutions.

We first consider the tax choice when all firms are making investment decisions ($I_{t-1} = 0$). Figure 4.3 plots the ex post optimal tax, the optimal one-period tax, the optimal two-period tax, and the break-even one-period tax for each probability p_t of investing. As expected, the ex post optimal tax slopes down in p_t and the break-even one-period tax does slope upward over a range of p_t. Firms expect the one-period tax to fall if they invest and so require a high one-period tax if they are to build low-pollution plant. The optimal one-period tax therefore must either be set high enough to spur some investment or high enough to obtain optimal emissions from conventional plant. In contrast, firms will invest for a lower two-period tax, though in comparison to its ex post optimal level, that two-period tax will be too high for two periods in a row. With the low investment cost, the optimal one-period tax induces some firms to invest, reducing the appeal of the two-period tax, but the regulator prefers the two-period tax for the higher investment cost. The two-period tax enables it to lower the tax from the optimal one-period level while inducing more firms to invest.



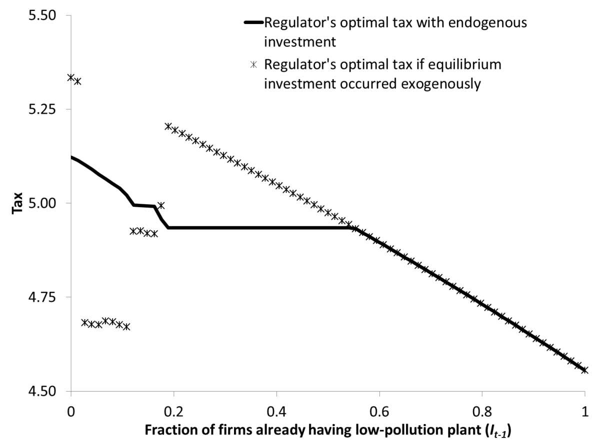
(a) Investment cost: 1.22 (b) Investment cost: 1.28

Figure 4.4: The equilibrium and socially optimal probability that firms with conventional plant invest in low-pollution plant as a function of the number of firms already having low-pollution plant. Also shows the probability that would leave the total level of low-pollution plant unchanged and the long-run level of low-pollution plant (calculated after investment) that the regulator soon obtains from any starting fraction of low-pollution plant.

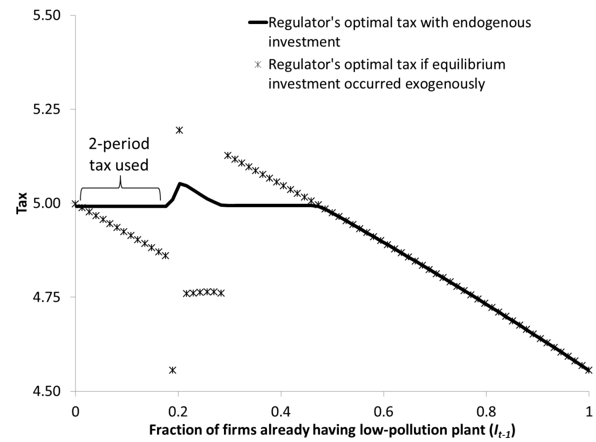
Does the multiperiod model produce socially optimal outcomes? Figure 4.4 compares equilibrium investment to socially optimal investment. Equilibrium investment is too high in both these cases when there is little low-pollution plant in the fleet, and the long-run fraction of low-pollution plant in the fleet is also greater than socially optimal in both cases. As we would expect from the analytic results, the suboptimally high investment also indicates a tax that would be set differently if the equilibrium investment would have occurred exogenously. When there is little low-pollution plant in the fleet, the regulator obtains investment with a tax that is greater than would be set if investment were taken as given (Figure 4.5). If it could lower the tax rate after firms invested but before they select emissions, it would. However, when there is already much low-pollution plant in the fleet, the regulator need not worry about investment incentives and so provides selects a tax that leads to optimal emissions. The multiperiod setting offers the regulator additional means of affecting lumpy investment by choosing the duration of this period's tax and, later, choosing the next period's tax rate, but both of these choices also affect emission decisions, which often prevents the equilibrium outcome from being socially optimal.

4.4 Discussion

We have shown that an ex ante linear emission tax often fails to achieve either socially optimal adoption of low-pollution green technology or socially optimal emissions. The fundamental



(a) Investment cost: 1.22



(b) Investment cost: 1.28

Figure 4.5: The regulator's optimal tax as a function of the number of firms already having low-pollution plant. Also shows the optimal tax if equilibrium investment (i.e., investment induced by the optimal tax) would have occurred exogenously. This tax provides optimal emission incentives conditional on equilibrium investment.

reason is that an emission tax fixes both investment and emission incentives in advance, and the regulator often cannot achieve both goals with this one instrument. In the one-period setting with identical investment costs, equilibrium investment outcomes are all-or-nothing, and the regulator often uses an ex post optimal tax that achieves too little investment or uses a tax set higher to stimulate investment. If firms are instead heterogeneous, only a fraction of firms may invest. Now, however, the regulator weights the tax between an ex post optimal tax and a tax that provides optimal investment incentives. Whether the resulting tax is above or below the ex post optimal tax depends on whether the marginal firm's investment costs change faster or slower than the tax rate in the neighborhood of the ex post optimal tax. Finally, in the multiperiod setting with identical investment costs, the longer lifetime of low-pollution plant makes the next period's anticipated tax affect the current period's investment, and the ability to select a two-period tax can enable additional investment. However, investment and emissions are still often suboptimal, and the cost of locking in the tax rate restricts the parameter space in which the regulator commits to the longer tax.

In all cases, we assume that a firm's investment is lumpy and takes effect immediately. If investment were not lumpy, then each firm's investment cost would increase with its quantity of investment. If firms could adjust investment continuously, then additional investment by each firm would have a more marginal effect on the tax rate. In this case, there is no externality from firms failing to account for the effect of investment on the optimal tax and the emission tax might produce the socially optimal outcome. If the effect of investment were delayed and the regulator could observe and react in the interim, then we would have a more standard hold-up problem in the analytic setting (Laffont and Tirole, 1996; Abrego and Perroni, 2002; Helm et al., 2004; Blyth et al., 2009). In the numeric setting, our results might not change overly much if the regulator must still always incentivize emissions and investment with one policy.

We have modeled the regulator as being locked into the chosen tax and, in the multiperiod setting, as being locked into the chosen duration. In reality, the regulator could change a tax at some cost. In addition to selecting the duration of a policy, the regulator also selects the cost of switching the policy. Modeling the interaction of endogenous switching costs and firms' investment opportunities could provide important insights for institutional design in climate policy. However, such a model would need to include stochasticity or learning on the part of the regulator. In our setting, the world changes as firms invest or not, but the regulator perfectly forecasts how it will change and whether it will be worth switching policy. The regulator would therefore always choose a switching cost of zero or infinity, which reduces to our multiperiod setting in which the regulator allows itself to either change policy costlessly or not at all. Endogenous switching costs become interesting when the regulator is not sure what the realized value of switching will be.

Our results imply that pollution policy should favor instruments that condition investment incentives on realized investment when obtaining adoption of low-pollution technology is a primary goal. In many pollution contexts, policy has dynamic goals that extend beyond controlling each year's emission to include developing new technologies, changing physical infrastructure, and inducing other investments that could affect the future cost of abatement. It is a well-known result that research and development (R&D) market failures call for a complementary policy to stimulate R&D (e.g., Fischer and Newell, 2008; Acemoglu et al., in press), and we here show that a linear tax could perform better if another instrument controlled the incentives for technology adoption. Yet specifying such a complementary instrument or a nonlinear tax often requires the regulator to observe investment. In contrast, tradable permits programs automatically condition the investment incentive on realized investment via the response of the permit price to the resulting shift in firms' demand for permits. Many factors are relevant to the choice of policy instrument (e.g., Weitzman, 1974; Requate, 2005; Hepburn, 2006), but the ability to use market mechanisms to condition the emission price on investment outcomes should join the list.

Finally, we have shown that it is not necessarily a priority to provide stable investment incentives unless there is an additional factor that would introduce instability (compare Blyth et al., 2007). In our multiperiod setting, investments are long-lived (as with much energy infrastructure), and policy can fluctuate in response to the changing composition of, for instance, electricity generation. However, the regulator rarely finds it optimal to provide more stable investment incentives by locking in the tax for longer periods. It may do so when not much infrastructure has been built, but it soon switches to a one-period tax once the first investments have been made. The foreseen fact that a tax will change both encourages use of a longer tax in order to maintain investment incentives but also discourages use of a longer tax because it distorts post-investment emission incentives. Policy stability would become more important with an outside source of instability such as changing information, changing politics, or changing preferences.

4.5 Appendix: Additional proofs

Proof of Proposition 1

The regulator chooses $s(I_0, p)$ to maximize firms' benefits net of the damages from pollution:

$$\begin{aligned}
&\max_s \Big\{ (1-p)(1-I_0)\pi_{noinv}(s) + p(1-I_0)\pi_{inv}(s) + I_0 \pi_{lowpoll}(s) \\
&\qquad - D\big(w_n e_n^*(s) + w_i e_i^*(s)\big) + s\left[w_n e_n^*(s) + w_i e_i^*(s)\right] \Big\} \\
&= \max_s \Big\{ w_n b_n(s) + w_i b_i(s) - p(1-I_0)c - D\big(w_n e_n^*(s) + w_i e_i^*(s)\big) \Big\} \qquad (4.13)
\end{aligned}$$

with p a parameter rather than a function of s. Note that each firm's first-order condition defines optimal emissions $e^*(s)$ as the level that equates marginal benefits to the chosen tax s: $b'(e^*) = s$. The regulator's first-order condition says that the ex post optimal tax should equate the marginal damage from emissions and the marginal benefit from emissions:

$$\begin{aligned}
&0 = w_n\ b_n'(e_n^*)\ e_n^{*\prime}(s) + w_i\ b_i'(e_i^*)\ e_i^{*\prime}(s) - \left[w_n e_n^{*\prime}(s) + w_i e_i^{*\prime}(s)\right]\ D'\big(w_n e_n^*(s) + w_i e_i^*(s)\big) \\
\Rightarrow\ &s = D'\big(w_n e_n^*(s) + w_i e_i^*(s)\big) \qquad (4.14)
\end{aligned}$$

where the second line substitutes s for $b_n'(e_n^*)$ and $b_i'(e_i^*)$ from the firms' first-order conditions.

Increasing the fraction of low-pollution plant initially in the fleet decreases the ex post optimal tax:

$$\begin{aligned}
\frac{\partial s}{\partial I_0} =& D''(x)\left[\frac{\partial w_n}{\partial I_0} e_n^*(s) + \frac{\partial w_i}{\partial I_0} e_i^*(s)\right] \\
=& \underbrace{D''(x)}_{>0}\, \underbrace{(1-p)}_{\geq 0}\, \underbrace{\left[e_i^*(s) - e_n^*(s)\right]}_{\leq 0}\ \leq 0 \qquad (4.15)
\end{aligned}$$

where $x \equiv w_n e_n^*(s) + w_i e_i^*(s)$. We know $e_n^*(s) \geq e_i^*(s)$ because they are each set to equate marginal benefit to the same tax s. Marginal benefit is weakly lower with low-pollution plant, so the low-pollution plant must emit less if its marginal benefit is to equal that of the conventional plant. Finally, increasing the fraction of plant that adopts low-pollution technology also decreases the ex post optimal tax:

$$\begin{aligned}
\frac{\partial s}{\partial p} =& D''(x)\left[\frac{\partial w_n}{\partial p} e_n^*(s) + \frac{\partial w_i}{\partial p} e_i^*(s)\right] \\
=& \underbrace{D''(x)}_{>0}\, \underbrace{(1-I_0)}_{\geq 0}\, \underbrace{\left[e_i^*(s) - e_n^*(s)\right]}_{\leq 0}\ \leq 0 \qquad (4.16)
\end{aligned}$$

The ex post optimal tax is lower when there is more low-pollution plant in the fleet because a lower tax can achieve the same emission level as with conventional plant. Having more low-pollution plant in the fleet increases the emission reductions produced by a given tax and so makes lower taxes look relatively more attractive.

Alternate proof of Proposition 6 using quadratic benefits and damages

We already saw that the regulator chooses to suppress investment with $\hat{\tau} - \epsilon$ if and only if $W_{w_i=I_0} > W_{w_i=1}$. Let benefits and damages be quadratic with $a_1 = 1$, and, for convenience, define $n \equiv I_0 + a_2(1 - I_0)$. The regulator chooses to suppress investment if and only if:

$$\begin{aligned} 0 > & n^2 \underbrace{\left\{ \frac{-d_2}{a_2 b_2} (\hat{\tau} - b_1)^2 \right\}}_{>0} + n \underbrace{\left\{ 2d_1(b_1 - \hat{\tau}) \right\}}_{>0} + \left\{ \underbrace{(\hat{\tau} + s)(\hat{\tau} - s)}_{<0} + \underbrace{2d_1(s - b_1)}_{<0} + \underbrace{\frac{d_2}{a_2 b_2}(s - b_1)^2}_{<0} \right\} \\ \equiv & K(n) \end{aligned} \tag{4.17}$$

The plausible range of I_0 (i.e., $I_0 \in [0, 1]$) implies that $n \in [1, a_2]$. We need to learn about the conditions under which $K(n) < 0$. Note that $K(0) < 0$ and $K(n)$ is increasing in n:

$$\frac{\partial K}{\partial n} = 2d_1(b_1 - \hat{\tau}) - \frac{d_2}{a_2 b_2} 2n(\hat{\tau} - b_1)^2 > 0 \tag{4.18}$$

Therefore the regulator suppresses investment for low n (high I_0) or not at all: if it doesn't suppress investment for I_0 close to 1 (i.e., for $n \approx 1$), then it never suppresses investment. The intuition is that while convex pollution damages nonlinearly increase the social benefits from investing as more firms invest, investment cost only increases linearly as more firms invest. Thus, $\exists\ I_0 \in [0, 1]$ such that the regulator suppresses investment if and only if $K(1) < 0$, and $K(1) < 0$ if and only if the following condition holds:

$$(\hat{\tau} + s)(d_2 - a_2 b_2) > 2(b_1 d_2 - d_1 a_2 b_2) \tag{4.19}$$

If we substitute for $s(I_0, 1)$ and for $\hat{\tau}$, this condition becomes:

$$\Delta_0 < \frac{-a_2 b_2 (b_1 - d_1)^2 (a_2 - 1)}{4(a_2 b_2 - d_2)^2} \equiv \phi \tag{4.20}$$

The regulator suppresses investment for some $I_0 \in [0, 1]$ if and only if $\Delta_0 < \phi$, and the regulator has the option of choosing such an equilibrium with suppressed investment if and only if $\Delta_0 \geq \psi$. However, note that $\phi = \psi$. Therefore if $\Delta_0 \geq \psi$ so that investment is an equilibrium with a tax of $s(I_0, 1)$, then the regulator will never choose to suppress investment for any plausible I_0.

Proof of Proposition 7

Define $s \equiv s(I_0, 0)$ and assume $s < \hat{\tau}$. The regulator compares welfare $W_{w_i=1}$ from the equilibrium with investment with welfare $W_{w_i=I_0}$ from the equilibrium without investment:

$$\begin{aligned} W_{w_i=I_0} =& (1-I_0)\pi_{noinv}(s) + I_0\pi_{lowpoll}(s) - D((1-I_0)e_n^*(s) + I_0 e_i^*(s)) \\ &+ s\left[(1-I_0)e_n^*(s) + I_0 e_i^*(s)\right] \end{aligned} \tag{4.21}$$

$$\begin{aligned} W_{w_i=1} =& (1-I_0)\pi_{inv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D(e_i^*(\hat{\tau})) + \hat{\tau}\ e_i^*(\hat{\tau}) \\ =& (1-I_0)\pi_{noinv}(\hat{\tau}) + I_0\pi_{lowpoll}(\hat{\tau}) - D(e_i^*(\hat{\tau})) + \hat{\tau}\ e_i^*(\hat{\tau}) \end{aligned} \tag{4.22}$$

It prefers to stimulate investment if and only if:

$$W_{w_i=1} > W_{w_i=I_0} \tag{4.23}$$

Let $a_1 = 1$ and define $n \equiv I_0 + a_2(1-I_0)$. The regulator chooses to stimulate investment if and only if:

$$\begin{aligned} \Rightarrow 0 >& n^2 \underbrace{(s-b_1)^2 \frac{d_2}{a_2 b_2}}_{<0} + n \underbrace{\left[2d_1(s-b_1) + \hat{\tau}^2 - s^2\right]}_{?} + \underbrace{\left\{2d_1(b_1-\hat{\tau}) - \frac{d_2}{a_2 b_2}(\hat{\tau}-b_1)^2\right\}}_{>0} \\ \equiv& J(n) \end{aligned} \tag{4.24}$$

Substituting for $s(I_0, 0)$ and $\hat{\tau}$, we find that the regulator stimulates investment if and only if:

$$0 > \underbrace{\frac{a_2 b_2 n (b_1-d_1)^2}{a_2 b_2 - d_2 n}}_{>0} + 4\left[\underbrace{(d_2 - a_2 b_2 n)\frac{\Delta_0}{a_2-1}}_{>0} + \underbrace{(d_1 - b_1 n)\sqrt{\frac{-a_2 b_2 \Delta_0}{a_2-1}}}_{<0}\right] \tag{4.25}$$

The sum in brackets is < 0. The regulator therefore stimulates investment if and only if:

$$\underbrace{\frac{-a_2 b_2 n (b_1-d_1)^2 (a_2-1)}{4(a_2 b_2 - d_2 n)}}_{<0} > \underbrace{(d_2 - a_2 b_2 n)\Delta_0}_{>0} + \underbrace{(d_1 - b_1 n)\sqrt{-a_2 b_2 (a_2-1)\Delta_0}}_{<0} \tag{4.26}$$

The first term on the right-hand side is positive while both the left-hand side and the second term on the right-hand side are negative. Therefore a necessary condition is that the rightmost term be less than the left-hand side, giving the first term in the following necessary condition for stimulating investment:

$$\max\left\{\frac{n^2(b_1-d_1)^2}{4(b_1 n - d_1)^2}\Psi,\ \Psi - \frac{2d_2 n (b_1-d_1)^2(a_2-1)}{4(a_2 b_2 - d_2 n)^2}\right\} < \Delta_0 \leq \Psi \tag{4.27}$$

The second term in the maximization operator is the condition from Proposition 8 for it being socially optimal for there to be no investment. The second inequality is the condition from Table 4.1 for the regulator's deciding between $\hat{\tau}+\epsilon$ and $s(I_0, 0)$.

Finally, note that $J(n)$ is globally concave in n for all Δ_0 and that, at $I_0 = 1$, $J(n; \Delta_0 = \Psi)$ is equal to 0. This means that when $\Delta_0 = \Psi$ and all plants are already low-pollution, the regulator is, reasonably, indifferent between stimulating investment or not because there are no plants that might invest. However, also observe that $J(n; \Delta_0 = \Psi)$ is increasing in I_0 (i.e., decreasing in n) at $I_0 = 1$. With the concavity of $J(n)$, this implies that $J(n; \Delta_0 = \Psi) < 1$ for all $n > 1$ (i.e., for all $I_0 < 1$). This in turn has two implications. First, for Δ_0 near Ψ and $I_0 \neq 1$, the regulator stimulates investment. Second, for Δ_0 near Ψ, the range of I_0 for which the regulator stimulates investment is a connected set bounded below by the allowed values of I_0 (i.e., by $I_0 = 0$). In order for this range of I_0 not to be a connected set, the sign of $J'(1)$ would have to be positive for the given Δ_0, which is not the case for Δ_0 sufficiently close to Ψ.

Proof of Proposition 8

The social optimum is found by a regulator who can choose both emissions and investment. If the regulator can choose investment, then the social optimum is obtained through a tax chosen with investment taken as exogenous, for firms face no additional externalities. We can therefore find the social optimum by finding the optimal investment probability and optimal tax conditional on investment without considering whether firms would actually invest in response to that tax:

$$\begin{aligned}
\max_{p,s} \Bigg\{ &\int_0^{I_0} \pi_{lowpoll}(s)\ dx + \int_{I_0}^{I_0+[1-I_0]p} \pi_{inv}(s)\ dx + \int_{I_0+[1-I_0]p}^{1} \pi_{noinv}(s)\ dx \\
&- D\left(\int_0^{I_0+[1-I_0]p} e_i^*(s)\ dx + \int_{I_0+[1-I_0]p}^{1} e_n^*(s)\ dx\right) \\
&+ s\left[\int_0^{I_0+[1-I_0]p} e_i^*(s)\ dx + \int_{I_0+[1-I_0]p}^{1} e_n^*(s)\ dx\right]\Bigg\} \\
&s.t.\ p \in [0,1],\ s \geq 0
\end{aligned} \tag{4.28}$$

We already solved the first-order condition for s in Proposition 1, so we can now just solve the first-order condition for p and substitute the solution for s. Assuming $a_1 = 1$, interior solutions give the socially optimal probability of investment p^* as that which solves:

$$\Delta_0 = \Psi + \frac{2d_2(b_1 - d_1)^2(a_2 - 1)}{4(a_2 b_2 - d_2 n)^2}[-I_0 + (1 - I_0)(-a_2 + 2p^*(a_2 - 1)] \equiv \chi(p^*) \tag{4.29}$$

$\chi(p^*)$ is monotonically increasing in p^*, meaning that greater investment cost decreases socially optimal investment. This also means that values of $p^* \in (0, 1)$ form a connected set in Δ_0. Values of Δ_0 greater than $\chi(1)$ or less than $\chi(0)$ take on the nearest corner solution for investment. When $\Delta_0 = \Psi$ and $I_0 < 1$, we have:

$$p^*_{\Delta_0=\Psi} = \frac{I_0 + a_2(1 - I_0)}{2(1 - I_0)(a_2 - 1)} > 0 \tag{4.30}$$

It is therefore socially optimal to invest with positive probability when $\Delta_0 = \Psi$.

Finally, we seek the condition under which it is socially optimal for all firms to invest even when the regulator's plausible ex post optimal tax would not produce any investment. This condition occurs when $p^*_{\Delta_0=\Psi} \geq 1$, which is equivalent to $I_0 \geq (a_2 - 2)/(a_2 - 1)$. Therefore, when $a_2 < 2$ or I_0 is sufficiently large, it is socially optimal for all firms to invest in some situations unattainable with the ex post optimal tax $s(I_0, 1)$.

Proof of Corollary 9

If $I_0 > 0$, we are done. Assume $I_0 = 0$. We are looking for the condition under which $\chi(0; I_0 = 0) \leq 0$, because in that case it is socially optimal for some firms to invest for all $\Delta_0 > 0$. The sign of $\chi(0; I_0 = 0)$ is the same as the sign of $-(b_2 + 2d_2)$, which gives us our result.

Proof of Proposition 10

Firms differ only according to their investment cost. Alternately, each firm has access to one low-pollution technology at the same investment cost, and while each technology has the same effect on the benefit from emitting, technologies differ according to their cost of adoption. The distribution of investment costs for firms making an investment decision is atomless and has cumulative distribution function $F(c)$ with support in the positive real numbers. Lemma 2 still applies to each firm, meaning that we can translate investment costs into break-even taxes according to some function $\hat{\tau} = g(c)$. Note that $g(c)$ is weakly increasing in investment cost and so is invertible. For any tax τ and fraction $1 - I_0$ of firms deciding whether to invest, the fraction of those firms that does invest is given by the cumulative distribution function $H(\tau)$, defined as $H(\tau) = F\left(g^{-1}(\tau)\right)$. When the regulator selects a tax, it knows the fraction of firms that will invest, and, in contrast to the model with identical investment costs, this fraction changes smoothly with the tax rate.

The regulator's problem is to choose the tax τ to maximize net social benefit, recognizing that the tax rate affects both emissions and the fraction of firms having low-pollution plant:

$$
\begin{aligned}
&\max_{\tau}\Bigg\{\int_0^{I_0}\pi_{lowpoll}(\tau)dx+\int_{I_0}^{1}\left[\int_0^{g^{-1}(\tau)}\pi_{inv}(\tau,c)\ F'(c)\ dc+\int_{g^{-1}(\tau)}^{\infty}\pi_{noinv}(\tau)F'(c)\ dc\right]dx\\
&\qquad-D\left(\int_0^{I_0+(1-I_0)H(\tau)}e_i^*(\tau)+\int_{I_0+(1-I_0)H(\tau)}^{1}e_n^*(\tau)\right)\\
&\qquad+\tau\left[\int_0^{I_0+(1-I_0)H(\tau)}e_i^*(\tau)+\int_{I_0+(1-I_0)H(\tau)}^{1}e_n^*(\tau)\right]\Bigg\}\\
&=\max_{\tau}\Bigg\{\int_0^{I_0+(1-I_0)H(\tau)}b_i(\tau)dx+\int_{I_0+(1-I_0)H(\tau)}^{1}b_n(\tau)dx-\int_{I_0}^{1}\int_0^{g^{-1}(\tau)}c\ F'(c)\ dc\\
&\qquad-D\left(\int_0^{I_0+(1-I_0)H(\tau)}e_i^*(\tau)+\int_{I_0+(1-I_0)H(\tau)}^{1}e_n^*(\tau)\right)\Bigg\}
\end{aligned}
\tag{4.31}
$$

Each firm's first-order condition still defines optimal emissions $e^*(\tau)$ as the level that sets its plant's marginal benefits to equal the tax τ: $b'(e^*)=\tau$. The ex post optimal tax $s(I_0,\tau)$ is analogous to the tax defined in Proposition 1, except now it is a function of the chosen ex ante tax τ instead of the exogenous probability p of investing:

$$
s(I_0,\tau)=D'\left(\int_0^{I_0+(1-I_0)H(\tau)}e_i^*(\tau)+\int_{(I_0)+(1-I_0)H(\tau)}^{1}e_n^*(\tau)\right)
\tag{4.32}
$$

The chosen ex ante tax determines the fraction of firms that invests. The regulator's first-order condition is:

$$
\begin{aligned}
0=&[I_0+(1-I_0)H(\tau)]\ b_i'(e_i^*)\ {e_i^*}'(\tau)+(1-I_0)\ H'(\tau)\ b_i(e_i^*(\tau))\\
&+[1-I_0][1-H(\tau)]\ b_n'(e_n^*)\ {e_n^*}'(\tau)-(1-I_0)\ H'(\tau)\ b_n(e_n^*(\tau))\\
&-(1-I_0)\ H'(\tau)\ g^{-1}(\tau)\ {g^{-1}}'(\tau)\\
&-[[I_0+(1-I_0)H(\tau)]\ {e_i^*}'(\tau)+(1-I_0)\ H'(\tau)\ e_i^*(\tau)+[1-I_0][1-H(\tau)]\ {e_n^*}'(\tau)-(1-I_0)\ H'(\\
&\qquad D'\left(\int_0^{I_0+(1-I_0)H(\tau)}e_i^*(\tau)+\int_{(I_0)+(1-I_0)H(\tau)}^{1}e_n^*(\tau)\right)\\
=&\underbrace{[w_i(\tau)\ {e_i^*}'(\tau)+[1-w_i(\tau)]\ {e_n^*}'(\tau)]\ \big[\tau-D'\big(x(\tau)\big)\big]}_{V_e(\tau)}\\
&+w_i'(\tau)\underbrace{\Big\{\underbrace{b_i(e_i^*(\tau))-b_n(e_n^*(\tau))-g^{-1}(\tau)\ {g^{-1}}'(\tau)}_{V_\pi(\tau)}+\underbrace{[e_n^*(\tau)-e_i^*(\tau)]\ D'\big(x(\tau)\big)}_{V_D(\tau)}\Big\}}_{V_I(\tau)}
\end{aligned}
$$

$$
\Rightarrow V_e(\tau^*)=-\ w_i'(\tau^*)\ \ \{V_\pi(\tau^*)+V_D(\tau^*)\}=-w_i'(\tau^*)\ V_I(\tau^*)
\tag{4.33}
$$

where the second equality substitutes τ for $b_i'(e_i^*)$ and $b_n'(e_n^*)$ from each firm's first-order condition, substitutes $x(\tau) \equiv \int_0^{I_0+(1-I_0)H(\tau)} e_i^*(\tau) + \int_{(I_0)+(1-I_0)H(\tau)}^1 e_n^*(\tau)$, and substitutes $w_i(\tau) \equiv [I_0 + (1 - I_0)H(\tau)]$ as the fraction of low-pollution plant after investment has occurred. This condition is composed of a few terms. The first is $V_e(\tau)$, or the value of the effect of a unit change in tax on emissions when investment is held constant. The condition $V_e(s) = 0$ defines the ex post optimal tax $s(I_0, \tau)$. The second term, $V_I(\tau)$, is the value from the effect of a unit change in tax on the fraction of firms having low-pollution plant. This term is itself composed of the change in the fraction of low-pollution plant, or $w_i'(\tau)$, multiplied by the sum of the effect of increasing investment on firms' pre-tax profits $V_\pi(\tau)$ and on damages $V_D(\tau)$.

The optimal tax considers efficiency in investment and in emissions. It is set so that the effect of a small tax change on fixed-investment emissions is offset by the effect on investment. If $w_i'(\tau) = 0$, then the optimal tax is the same as the ex post optimal tax because changing the tax does not affect investment. This occurs when $I_0 = 1$ or if $H'(\tau) = 0$. It is always the case that $V_D(\tau) \geq 0$ and $V_\pi(\tau) \leq 0$, so the sign of the right-hand side in equation (4.33) depends on whether investment has a stronger effect on pre-tax profits or on damages. If the effect on damages is stronger, the optimal tax will be above marginal damage so that $V_e(\tau^*) < 0$. If the effect on pre-tax profits is stronger, then the optimal tax will be set below marginal damage so that $V_e(\tau^*) > 0$. If the tax were set to the level s that has $V_e(s) = 0$, then V_I becomes:

$$\begin{aligned} V_I(s) =& b_i(e_i^*(s)) - b_n(e_n^*(s)) - g^{-1}(s)\ g^{-1\prime}(s) + s\ [e_n^*(s) - e_i^*(s)] \\ =& \pi_{inv}(s, g^{-1}(s)\ g^{-1\prime}(s)) - \pi_{noinv}(s) \end{aligned} \tag{4.34}$$

This additional pre-tax profit due to the next firm's investment is equal to 0 when $g^{-1\prime}(s) = 1$, in which case the ex post optimal tax also provides optimal investment incentives. However, if $g^{-1\prime}(s) < 1$, then the marginal firm's investment cost does not increase as fast as the tax rate and the ex post optimal tax obtains inefficiently low investment. If $g^{-1\prime}(s) > 1$, then the marginal firm's investment cost increases faster than the tax rate, and the ex post optimal tax provides too much incentive to invest. Therefore, the optimal tax τ^* is greater than the ex post optimal tax $s(I_0, \tau^*)$ if the marginal firm's investment cost responds weakly to a change in the tax rate from s. This is because a higher tax rate can obtain more investment at relatively low additional cost. However, the optimal tax τ^* is less than the ex post optimal tax $s(I_0, \tau^*)$ if the marginal firm's investment cost responds strongly to a change in the tax rate from s. In this case, increasing the tax rate obtains more investment at relatively high additional cost. In general, the ex post optimal tax need not produce socially optimal investment because it does not account for the distribution of investment costs. It is set based on emissions' response to a tax change without considering how emissions also respond to firms' altered investments. Because the fraction of firms investing responds smoothly to the chosen tax, the regulator does not face multiple equilibria as in the case with homogeneous firms, but it will often not choose the ex post optimal tax when setting the tax ex ante because it still must worry about investment incentives.

Chapter 5

The influence of negative emission technologies and technology policies on the optimal climate mitigation portfolio[1]

Combining policies to remove carbon dioxide (CO_2) from the atmosphere with policies to reduce emissions could decrease CO_2 concentrations faster than possible via natural processes. We model the optimal selection of a dynamic portfolio of abatement, research and development (R&D), and negative emission policies under an exogenous CO_2 constraint and with stochastic technological change. We find that near-term abatement is not sensitive to the availability of R&D policies, but the anticipated availability of negative emission strategies can reduce the near-term abatement optimally undertaken to meet 2°C temperature limits. Further, planning to deploy negative emission technologies shifts optimal R&D funding from "carbon-free" technologies into "emission intensity" technologies. Making negative emission strategies available enables an 80% reduction in the cost of keeping year 2100 CO_2 concentrations near their current level. However, negative emission strategies are less important if the possibility of tipping points rules out using late-century net negative emissions to temporarily overshoot the CO_2 constraint earlier in the century.

[1]Thanks to Dan Kammen and to my co-authors at IIASA: Sabine Fuss, Jana Szolgayova, and Michael Obersteiner.

5.1 Introduction

International agreements advocate limiting global average temperature change to 2°C or less (MEF, 2009; UNFCCC, 2009). Because the world may have already used up half of the resulting carbon budget since 1750 (Allen et al., 2009), staying within the remaining carbon budget would prove challenging even with aggressive near-term abatement (e.g., van Vuuren et al., 2007). This challenge is further exacerbated as countries delay abatement (Clarke et al., 2009; Krey and Riahi, 2009). The dissonance between climate goals and action has spurred recent interest in additional ways of managing temperature outcomes (e.g., Keith, 2009; Lenton and Vaughan, 2009; Blackstock and Long, 2010; Kintisch, 2010). First, geoengineering techniques might reduce the temperature increase resulting from a CO_2 emission path by, for instance, reflecting more incoming solar radiation back into space. Second, large-scale use of negative emission technologies (NETs) can remove previously emitted atmospheric CO_2 and make an emission path partially reversible.[2] One type of NET is an air capture facility that directly removes CO_2 from ambient air via chemical reactions (e.g., Stolaroff et al., 2008). These technologies are still under development and their cost is uncertain (Keith, 2009). A second type of NET combines carbon capture and storage (CCS) technology with biomass-fired electricity generation (e.g., Rhodes and Keith, 2005; Uddin and Barreto, 2007). CCS is often discussed as a means of reducing the CO_2 emissions from coal-fired power plants, but it can also be used to capture the CO_2 that biomass previously absorbed from the atmosphere. In many modeling studies, bioenergy with carbon capture and storage (BECCS) makes low CO_2 concentration targets possible by turning the energy sector into a net carbon sink (e.g., Fisher et al., 2007; Clarke et al., 2009; Azar et al., 2010; Edenhofer et al., 2010; van Vuuren et al., 2010b).

We model climate policy portfolios with options to reduce emissions, to directly fund research and development (R&D) into low-carbon technologies, and to deploy NETs. The goal is to assess how the presence of different policy options affects optimal emission paths and policy costs. Most previous analyses of optimal policy portfolios have not included negative emission options (e.g., Fischer and Newell, 2008; Gerlagh et al., 2009), and analyses that considered NETs did not embed them in a setting with R&D options. Among these, Keith et al. (2006) used an integrated assessment model to explore how possible air capture of CO_2 affects climate strategies motivated by the possibility of abrupt climate change. They found that the future availability of air capture could reduce near-term abatement efforts but increase net long-term abatement, potentially returning atmospheric CO_2 concentrations to

[2] The captured CO_2 would be moved to geological sequestration absent another use or form of storage (e.g., Stephens and Keith, 2008). Importantly, geological sequestration of CO_2 can pose its own risks, and leakage can reduce the effectiveness of negative emission technologies (Benson et al., 2005; Damen et al., 2006; van der Zwaan and Gerlagh, 2009). Other negative emission strategies include methods that use biological activity to sequester atmospheric CO_2 (Read, 2009; Woodward et al., 2009), such as applying biochar to soils (Lehmann, 2007), sending crop residues to the deep ocean (Strand and Benford, 2009), and fertilizing swathes of ocean to promote plankton blooms (Smetacek and Naqvi, 2008; Strong et al., 2009).

pre-industrial levels within 200 years. Azar et al. (2006) and Azar et al. (2010) found that bioenergy with carbon capture and storage can be quite valuable in enabling more ambitious CO_2 targets (such as 350 ppm) but is less valuable if CO_2 targets are closer to 450 ppm. Our model has less technological detail but more policy options, thereby providing insight into how NETs may influence climate policy portfolios.

In addition to including NETs, we extend previous literature on the interaction between optimal abatement and R&D policies in the presence of endogenous technological change. Goulder and Mathai (2000) explored the implications of possible R&D investments and of learning-by-doing on optimal carbon taxes and abatement in both a cost-effectiveness and a benefit-cost setting.[3] Recognizing that both processes for producing technological change are important (Clarke et al., 2006), we include both channels in a cost-effectiveness setting: technological change can occur through public R&D policies, and technological change can also occur through the influence of abatement policies via learning-by-doing and private R&D.[4] Because it is important to explicitly model uncertainty when evaluating the optimal strength of technology policy (Baker and Shittu, 2008), we make technological change stochastic by adapting a three-point probability distribution (Baker and Adu-Bonnah, 2008). Further, we model two types of technological progress: one that is more valuable at lower levels of abatement, and one that is more valuable at higher levels of abatement (Baker and Adu-Bonnah, 2008). Technological change can therefore have two different types of impacts on abatement cost, and the realization of each type of technological change depends stochastically on both public R&D and abatement.

We combine these technology policy options in a single stylized numerical model that also includes options to reduce emissions and to research and deploy NETs. We do not predict optimal policy paths but instead seek robust insights from a framework complex enough to have multiple interacting policy options. We explore how these policy options influence the portfolio that meets an exogenous CO_2 constraint at the least expected cost. The CO_2 constraint is fixed and known in a given model run, but technological change depends stochastically on previous abatement and R&D funding and policy choices can respond to observed technological change. We next describe the model for optimally selecting a climate policy portfolio in each of three periods over the 21st century. We then present the results of solving it with stochastic programming for several parameterizations and constraints. The results illustrate the implications of future negative emission options for optimal near-term abatement and R&D efforts and for the cost of policy portfolios. They also demonstrate how making policy avoid threshold effects from temporarily high CO_2 levels affects the value and timing of NET deployment.

---

[3]While Goulder and Mathai (2000) used induced technological change (ITC) to refer to the effect on future abatement technology of both abatement and direct public R&D support, we reserve ITC to refer only to the effect of abatement. We do not, however, restrict abatement to only affect technology via learning-by-doing.

[4]Our two channels are similar to the two-factor experience curves summarized by Clarke et al. (2008). We do not consider how knowledge spillovers might affect the balance between R&D and abatement policies in the presence of induced technological change (see Hart, 2008; Greaker and Pade, 2009).

5.2 Model of policy portfolio optimization

We model a global decision-maker planning abatement, R&D funding, and NET deployment over the 21st century. Combining several types of policy options in one model enables interactions that might not be apparent otherwise. The decision-maker selects the most cost-effective policy plan for meeting a predetermined emission goal. These policy plans are contingent on low-carbon technological outcomes drawn from probability distributions determined by R&D funding and by abatement. In reality, global climate policy emerges from a game played among many decision-makers with complex objectives, but the case with a single decision-maker provides one benchmark for establishing and assessing climate policies.[5] In order to gain intuition for how these policy options interact, we compare results for three emission constraints and for four worlds: one world in which the only policy option is to reduce emissions; one world in which abatement and public R&D are both available options; one world in which abatement, public R&D, and NET research and deployment are all available options; and a final world in which all these options are available but temporarily overshooting the emission constraint is not allowed.

The objective is to select the dynamic policy portfolio that minimizes the cost of meeting an exogenous constraint e^* on cumulative CO_2 emissions. The policy levers available to the decision-maker are different levels of abatement $\{\mu_t\}_{t=1}^3$, of NET deployment $\{\kappa_t\}_{t=1}^3$, of carbon-free public R&D $\{\hat{\alpha}_t\}_{t=1}^3$, of emission intensity public R&D $\{\hat{\gamma}_t\}_{t=1}^3$, and of NET public R&D $\{\hat{\phi}_t\}_{t=1}^3$ (Table 5.1):

$$\begin{aligned}
&\min_{\mu_1,\kappa_1,\hat{\alpha}_1,\hat{\gamma}_1,\hat{\phi}_1} \Bigg\{ \mu_1 e_1 c(\mu_1,\alpha_1,\gamma_1) + f(\kappa_1,\phi_1) + g\left(\frac{\hat{\alpha}_1}{\bar{\alpha}}\right) + h\left(\frac{\hat{\gamma}_1}{\bar{\gamma}}\right) + j\left(\frac{\hat{\phi}_1}{\bar{\phi}}\right) \\
&+ \beta^{20} E_1\Bigg[\min_{\mu_2,\kappa_2,\hat{\alpha}_2,\hat{\gamma}_2,\hat{\phi}_2} \Big\{ \mu_2 e_2 c(\mu_2,\alpha_2,\gamma_2) + f(\kappa_2,\phi_2) + g\left(\frac{\hat{\alpha}_2}{\bar{\alpha}}\right) + h\left(\frac{\hat{\gamma}_2}{\bar{\gamma}}\right) + j\left(\frac{\hat{\phi}_2}{\bar{\phi}}\right) \\
&+ \beta^{20} E_2\Big[\min_{\mu_3,\kappa_3} \{\mu_3 e_3 c(\mu_3,\alpha_3,\gamma_3) + f(\kappa_3,\phi_3)\}\Big]\Big\}\Bigg]\Bigg\}
\end{aligned} \tag{5.1}$$

[5] In a model with R&D, having multiple actors introduces the possibility of international spillovers, which tend to reduce equilibrium R&D investment (Bosetti et al., 2008). NETs might have complex effects with multiple actors. On the one hand, NETs introduce a means for one country to unilaterally take care of another's emissions, but on the other hand, they also increase the scope for free-riding on others' emission reductions.

Table 5.1: Key to notation for decision variables and important parameters

Symbol	Description
Decision variables	
μ_t	Abatement at time t
κ_t	Negative emission technology (NET) deployment at time t in Gt CO_2
$\hat{\alpha}_t$	Target for time $t+1$ technology selected by time t public R&D into carbon-free technologies
$\hat{\gamma}_t$	Target for time $t+1$ technology selected by time t public R&D into emission intensity technologies
$\hat{\phi}_t$	Target for time $t+1$ NET cost selected by time t public R&D into NETs
Parameters	
e_t	Business-as-usual (BAU) emissions at time t (Gt CO_2)
e^*	Maximum cumulative emissions over the century (Gt CO_2)
S	Set of periods in which the cumulative emission constraint applies, which determines whether temporary overshoots are allowed
ν_α, ν_γ	Effectiveness of abatement at inducing technological change
$\bar{\alpha}, \bar{\gamma}, \bar{\phi}$	Maximal possible technological advance
$\alpha_t, \gamma_t, \phi_t$	Realized time t technology outcomes
$p_\alpha, p_\gamma, p_\phi$	Probability of missing the technology target implied by R&D and abatement

$$\text{subject to} \quad \sum_{t=1}^{s}(1-\mu_t)e_t - \kappa_t \leq e^*, \, \forall s \in S \tag{5.2}$$

$$\begin{aligned} \text{Transition probabilities:} \quad & \alpha_{t+1}\text{: see equations (5.3) through (5.5)} \\ & \gamma_{t+1}\text{: see equations (5.6) through (5.8)} \\ & \phi_{t+1}\text{: see equations (5.9) through (5.11)} \end{aligned}$$

Time t expectations E_t are over time $t+1$ technology outcomes and depend on time t technology, time t public R&D funding, and time t abatement (see appendix). In each period, the decision-maker observes the current technology and optimizes accordingly. The periods correspond to 2010-2029, 2030-2049, and 2050-2099, which roughly match the near-term, intermediate-term, and long-term periods for which CO_2 emission goals are often discussed. Scenarios vary the planner's access to certain types of policies by varying the possible levels that each decision variable may take (Table 5.2). μ_t gives the fraction of business-as-usual (BAU) emissions e_t abated in period t, and κ_t gives the quantity (Gt CO_2) of NETs deployed. Carbon-free technology reduces abatement cost by a fraction α_t, and emission intensity technology reduces non-abated emissions by a fraction γ_t (Baker and Adu-Bonnah, 2008). Carbon-free technology is relatively more valuable at high levels of abatement when abatement cost is correspondingly high. As an example, consider battery and renewable generation breakthroughs for all-electric vehicles. Emission intensity technology is relatively more valuable at lower levels of abatement when there are more non-abated emissions. As an example, consider powertrain technology that enables gasoline-electric hybrid vehicles. R&D into NETs can reduce the cost of deploying NETs by a fraction ϕ_t. The average cost of abatement ($c(\cdot)$) depends on the fraction of BAU emissions abated (μ_t) and on the outcomes of previous R&D into carbon-free technologies (α_t) and emission intensity technologies (γ_t). The cost of NETs ($f(\cdot)$) depends on the level of deployment (κ_t) and on the outcome of past R&D efforts (ϕ_t). R&D funding ($g(\cdot)$, $h(\cdot)$, and $j(\cdot)$) is determined by the chosen public R&D targets, and the total technology target for a period is determined by the public R&D target and by abatement policies' induced technological change (ITC, see appendix). The discount factor β converts costs from their value at the beginning of the period in which they are incurred to their value in the prior year.

Abatement, R&D, and NET deployment is motivated by the cumulative CO_2 emission constraint e^*. Cumulative emissions are a robust indicator of total temperature change (Allen et al., 2009; Matthews et al., 2009; National Research Council, 2011). As CO_2 concentrations are a conventional way of framing climate policy, we also convert each of these cumulative emission constraints to a year 2100 CO_2 concentration by assuming a constant airborne fraction of 0.45 (Denman et al., 2007). We model three values for e^* (Table 5.2): 88 Gt CO_2 (390 ppm), 880 Gt CO_2 (435 ppm), and 2900 Gt CO_2 (550 ppm). If there are no further CO_2 emissions, these cumulative emissions ultimately produce temperature change of 1°C, 2.5°C, and 6°C, respectively, under the best estimate from National Research Council (2011). If, instead, emissions continue past 2100 at a level that stabilizes the CO_2 concen-

Table 5.2: Policy options available in each scenario. Also, the three constraints on cumulative CO_2 emissions combined with each policy option scenario. See Table 5.1 for a key to the notation.

	Decision variables			Parameters	
	$\{\mu\}_{t=1}^{3}$	$\{\kappa\}_{t=1}^{3}$	$\{\hat{\alpha},\hat{\gamma},\hat{\phi}\}_{t=1}^{3}$ [a]	S	e^{*} [b]
Policy environment					
Only abatement	$\left\{0,\frac{1}{4},\frac{1}{2},\frac{3}{4},1\right\}$	0	0	$\{3\}$	
+R&D [c]	$\left\{0,\frac{1}{4},\frac{1}{2},\frac{3}{4},1\right\}$	0	$\left\{0,\frac{\bar{y}}{4},\frac{\bar{y}}{2},\frac{3\bar{y}}{4},\bar{y}\right\}$	$\{3\}$	
+R&D,NETs	$\left\{0,\frac{1}{4},\frac{1}{2},\frac{3}{4},1\right\}$	$\left\{0,\frac{e_3}{10},\frac{e_3}{4},\frac{e_3}{2},e_3\right\}$	$\left\{0,\frac{\bar{y}}{4},\frac{\bar{y}}{2},\frac{3\bar{y}}{4},\bar{y}\right\}$	$\{3\}$	
+R&D,NETs,Threshold [d]	$\left\{0,\frac{1}{4},\frac{1}{2},\frac{3}{4},1\right\}$	$\left\{0,\frac{e_3}{10},\frac{e_3}{4},\frac{e_3}{2},e_3\right\}$	$\left\{0,\frac{\bar{y}}{4},\frac{\bar{y}}{2},\frac{3\bar{y}}{4},\bar{y}\right\}$	$\{1,2,3\}$	
Constraint on year 2100 CO_2 [e]					
390 ppm					88
435 ppm					880
550 ppm					2900

[a] Values shown use y as a stand-in for the variable of interest. y should be replaced by α, γ, and ϕ as appropriate.
[b] Gt CO_2
[c] R&D to lower the cost of NETs is irrelevant when in a world in which NETs are unavailable.
[d] A climate threshold occurring at e^{*} rules out temporarily overshooting the cumulative emission constraint. This threshold is irrelevant in scenarios without available NETs because cumulative emissions cannot be reversed anyway.
[e] Implemented as a constraint on cumulative 21st century CO_2 emissions.

tration at its year 2100 value, then the 550 ppm constraint corresponds to requiring a 90% chance of keeping temperature change below 4°C, the 435 ppm constraint corresponds to requiring a 95% chance of keeping temperature change below 4°C, and the 390 ppm constraint corresponds to requiring a 90% chance of keeping temperature change below 2°C (Lemoine, 2010a). BAU emissions e_t (in Gt CO_2) come from scenario A2r in the International Institute for Applied System Analysis (IIASA) GGI Scenario Database (see also Riahi et al., 2007).[6] Summing over each period's years yields:

$$e_1 = 750,\ e_2 = 1150,\ e_3 = 4500$$

Assuming a constant airborne fraction of 0.45, the BAU path produces CO_2 concentrations of 428 ppm in 2030, 493 ppm in 2050, and 749 ppm in 2100.[7] All three modeled constraints could be met solely by abatement; NET deployment and public R&D funding are not necessary to meet the constraints.

We model two versions of each constraint on 21st century cumulative emissions (Table 5.2): one that allows temporary overshoots provided the constraint is met at the end of the century, and one that does not allow temporary overshoots during the century (compare Clarke et al., 2009). To the extent that 21st century temperature change is determined by 21st century cumulative emissions, temporary overshoots are consistent with the temperature limit that motivates constraining cumulative emissions. In this case, we have $S = \{3\}$. The freedom to temporarily overshoot the cumulative emission constraint only matters in a world with NETs, because cumulative emissions otherwise cannot decrease. However, using NETs to temporarily overshoot a cumulative emission constraint could cause additional irreversible changes or spur tipping points (O'Neill and Oppenheimer, 2004; Lenton et al., 2008). It is therefore also of interest to consider a world with NETs but where climate science indicates that a threshold would be crossed if cumulative emissions exceed e^*.[8] In this case, we have $S = \{1, 2, 3\}$, which requires the cumulative emission constraint to be met in each period rather than only in the final period.

The appendix describes the three-point probability distributions that determine the technology outcomes (α_t, γ_t, and ϕ_t) that apply to period t. It also describes how abatement induces technological change and defines the cost functions for abatement, NET deployment, and public R&D targets. Induced technological change here includes all private R&D and learning-by-doing that occur in response to an abatement policy. It does not include technological change due to public R&D policies, which are decision variables, or to spillovers,

[6] Available at: http://www.iiasa.ac.at/Research/GGI/DB/

[7] Experiments using the lower BAU emissions from scenario B2 showed that our results are robust to assumptions about the BAU path. The difference between BAU emission paths can represent different assumptions about population growth, the distribution of worldwide economic growth, future consumption habits, and BAU low-carbon technology adoption.

[8] In contrast to this firm constraint, Keller et al. (2004) and Lemoine and Traeger (2010) modeled tipping points as affecting the climate system or climate damages in a future world. They considered the decision about whether to risk crossing a possibly uncertain threshold, whereas we here take it as given that a policymaker has decided not to cross a known threshold.

which are not modeled (see Clarke et al., 2008). We solve the model by working backwards through the graph of all possible states. Each of the 15 parameterizations (see Table 5.3 in the appendix) is run under each of 9 combinations of the constraints on cumulative emissions and available policy options (Table 5.2). Each model run yields the optimal policy portfolio in each period conditional on previous technological outcomes and on previous abatement and NET policies. Comparing model runs reveals the importance of R&D and negative emission options, of the CO_2 constraint, and of other key parameters. It also refines intuition about how policy options interact. Because the parameterizations are not fully calibrated to empirical work, and because the correct process and distribution for technological change cannot be known in advance, the model's results should not be read for the recommended level of various policy variables. The goal is instead to assess the robustness of optimal portfolios and the crucial parameters for determining those portfolios.

5.3 Results: Portfolio cost, robust actions, and critical parameters

Tighter climate constraints require more expensive policy portfolios, but the relative cost of those portfolios depends strongly on the available policy options (Figure 5.1). R&D options provide their greatest cost reductions in percentage terms for weaker CO_2 constraints while NETs provide their greatest cost reductions for stricter constraints. R&D options provide their greatest percentage cost reductions for the weaker CO_2 constraints because these constraints permit greater flexibility in the timing of abatement and so allow abatement to be adjusted to take advantage of R&D outcomes. In the base case parameterization, including options to undertake R&D reduces the expected cost of meeting the 390 ppm constraint by almost 25%, reduces the expected cost of meeting the 435 ppm constraint by 55%, and reduces the expected cost of meeting the 550 ppm constraint by around 65%. In contrast, NETs provide their greatest expected cost reductions for the strictest CO_2 constraints because they then replace more expensive abatement. Including options to deploy NETs reduces the expected cost of the 390 ppm constraint by almost a further 80%, reduces the expected cost of the 435 ppm constraint by a further 35%, and does not further reduce the expected cost of the 550 ppm constraint. In the base case parameterization with NETs, the policy portfolio for the 390 ppm constraint costs about as much as the portfolio with R&D options for the 435 ppm constraint and about twice as much as the abatement-only portfolio for the 550 ppm constraint. However, when scientific findings lead policymakers to require that the emission constraint never be crossed, NETs provide less value because there is less flexibility to reallocate emissions over time.

Portfolio costs for the two more stringent constraints vary widely among parameterizations, as shown by the error bars in Figure 5.1. In all cases, the parameterizations that produce the most expensive policy portfolios are the parameterization without discounting

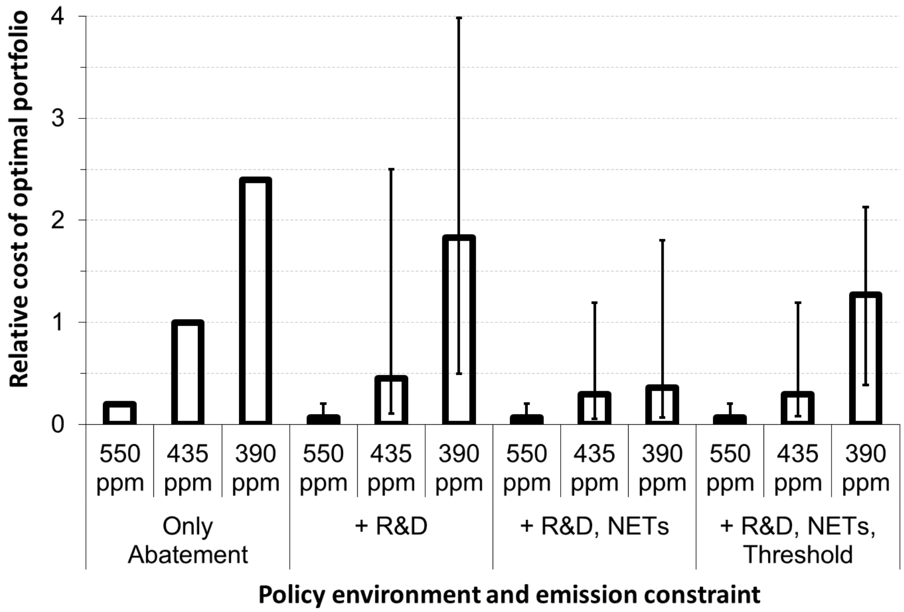


Figure 5.1: The present expected cost of the optimal policy portfolio in the base case scenarios. Error bars show the range across parameterizations. Costs are given as multiples of the cost in the 435 ppm scenario with abatement as the only policy option.

and the parameterization that limits the maximal scope of technological advance. The lack of discounting increases the present cost of late-century deep abatement, and limiting the scope of technological advance increases the expected cost of late-century abatement and NET deployment. In most cases, the two parameterizations that yield the lowest-cost policy portfolios are the parameterization with high discounting and the parameterization with low-cost abatement, low-cost R&D, and low-cost NETs. However, the most stringent (390 ppm) CO_2 constraint is a bit different. Here, if NETs are unavailable, the optimal portfolio in the parameterization with low-cost abatement is cheaper than the optimal portfolio in the parameterization with high discounting; if NETs are available and temporary overshoots are allowed, the optimal portfolio in the parameterization with high discounting is cheaper because so many of the costs result from the final period's heavy use of NETs; and if temporary overshoots are not allowed, the optimal portfolio in the parameterization with low-cost NETs is cheaper than the optimal portfolio in the parameterization with high discounting because NETs must be used in earlier periods.

The presence of R&D and NET options can affect not just the cost of the policy portfolio but also the optimal emission path. The lines with squares in Figure 5.2 show the optimal emission path if the only policy option is to undertake abatement. The lines with circles show the BAU emission path, which is scenario-independent. Each solid line represents the optimal gross emission path (i.e., before subtracting NETs' removed emissions) in the modeled parameterizations, with the thickness of a line proportional to the number of represented parameterizations. Comparing the solid lines to the one with squares shows how including a set of policy options changes the emission path relative to a case in which the only policy option is for abatement. Comparing solid lines across columns shows the effect on optimal emissions of including additional policy options or climate threshold constraints. Finally, comparing solid lines across rows shows the effect of the CO_2 constraint on optimal emissions.

Some have argued that technology policies should be the primary component of near-term climate policy (e.g., Sandén and Azar, 2005; Montgomery and Smith, 2007),. Our model would support this argument if making public R&D policies available shifted abatement from earlier periods to later ones. Instead, the left column in Figure 5.2 shows that this model's planned abatement paths are relatively insensitive to the availability of public R&D options, even though those options are exercised and do reduce portfolio costs. While Goulder and Mathai (2000) found that the availability of R&D options should affect optimal abatement, we find a small effect on abatement in part because we limit abatement to five discrete levels. R&D's expected effect on the cost of future abatement is high enough to justify undertaking it, but it is not high enough to reshuffle the intertemporal allocation of abatement between these five levels. In contrast, comparing the left column with the middle column shows that NET options do affect optimal emission paths: with the 435 ppm CO_2 constraint (middle row), making NETs available allows more smoothing of emissions over time by offsetting the most expensive late-century abatement, and with the 390 ppm CO_2 constraint (bottom row), NETs' availability decreases both near-term and long-term abatement by enabling future NET deployment to offset increased emissions from earlier periods.[9] The graphs for the 550 ppm constraint do not change, reflecting the insensitivity of emissions to NET options with a lax CO_2 constraint. Finally, comparing the right column with the middle column shows the influence of concerns about climate tipping points on optimal emission paths. Now the scenarios with the 390 ppm constraint (bottom row) increase both abatement and NET deployment in the first period so that CO_2 concentrations do not temporarily overshoot the target value. Therefore, while NET options can reduce optimal near-term abatement, the magnitude of this effect is sensitive to whether emissions and CO_2 concentrations may temporarily overshoot their year 2100 constraints.

In a stylized model such as the present one, the details of the control variables are less important than the big-picture story they represent. In Figure 5.3, we group cost-minimizing policy outcomes according to the probability with which they produce at least 25% abatement in the first period, at least 50% abatement in the second period, 100% abatement in the third period, public funding for carbon-free R&D in any period, public funding for emission intensity R&D in any period, and deployment of NETs in any period.

[9] The quantities of NETs deployed are within the range of estimates of underground global CO_2 storage capacity (Benson et al., 2005). While NETs might not involve underground storage, captured CO_2 from fossil fuel plants could also compete with captured CO_2 from negative emission facilities for end uses or storage capacity.

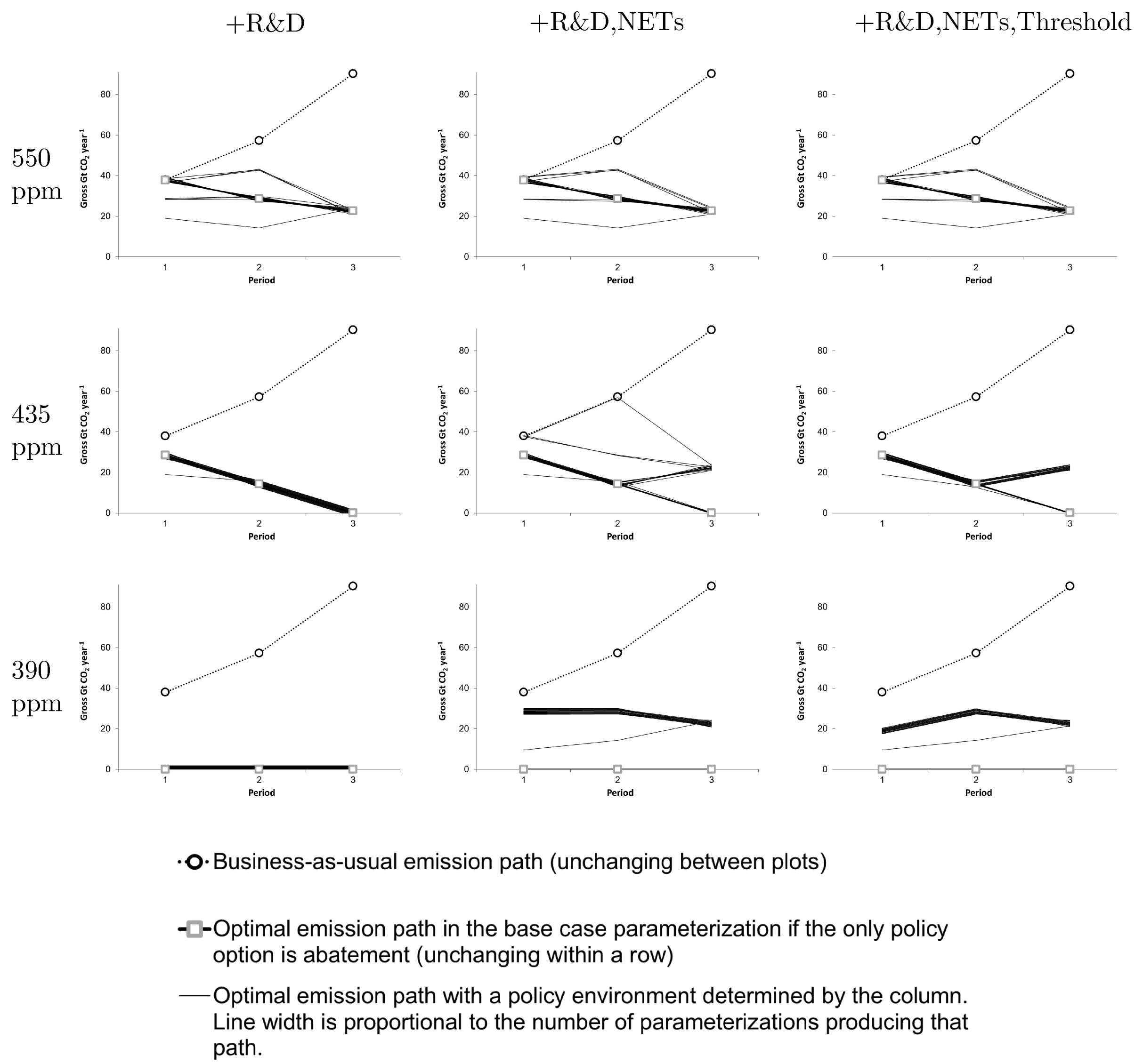


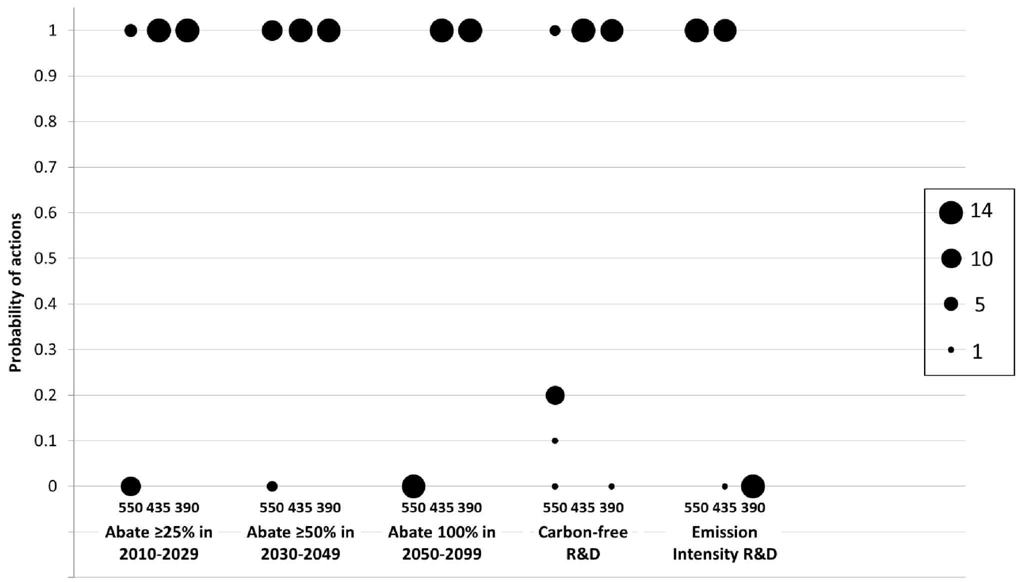
Figure 5.2: The planned gross emission paths (before subtracting NETs' removed CO_2) under the three year 2100 CO_2 constraints (rows) with different sets of available policy options (columns) (Table 5.2). Each chart shows the business-as-usual path (circles) and the base case planned path if the only available options are for abatement (squares). Each solid line represents the planned actions in the presence of options beyond abatement, where a planned action is the most likely action conditional on the previous most likely actions.

A probability strictly between 0 and 1 reflects that optimal policies can depend on realizations of stochastic technology. Interestingly, the probability of undertaking these broad categories of actions generally splits into probabilities near 1 and near 0. This indicates that big-picture actions are not sensitive to technological outcomes, instead being driven mostly by the CO_2 constraint. The type of R&D funded depends on how much it may contribute to the broad categories of actions favored by a given combination of CO_2 constraint and available policy options: carbon-free public R&D and emission intensity public R&D often substitute for each other, with expectations of future abatement largely driving the choice between the two types of technology forcing. In a subtle difference from the conclusions of Gerlagh et al. (2009) and of the review by Baker and Shittu (2008), near-term abatement and public R&D funding do not clearly substitute for each other with a given emission constraint and our discretized policy levels. Rather, near-term abatement is primarily determined by whether it is needed to keep future CO_2 concentrations below the constraint, not by the availability of R&D policies, which in turn are adopted without reducing near-term abatement. Near-term abatement is affected more by the availability of NETs than by the availability of R&D policies.

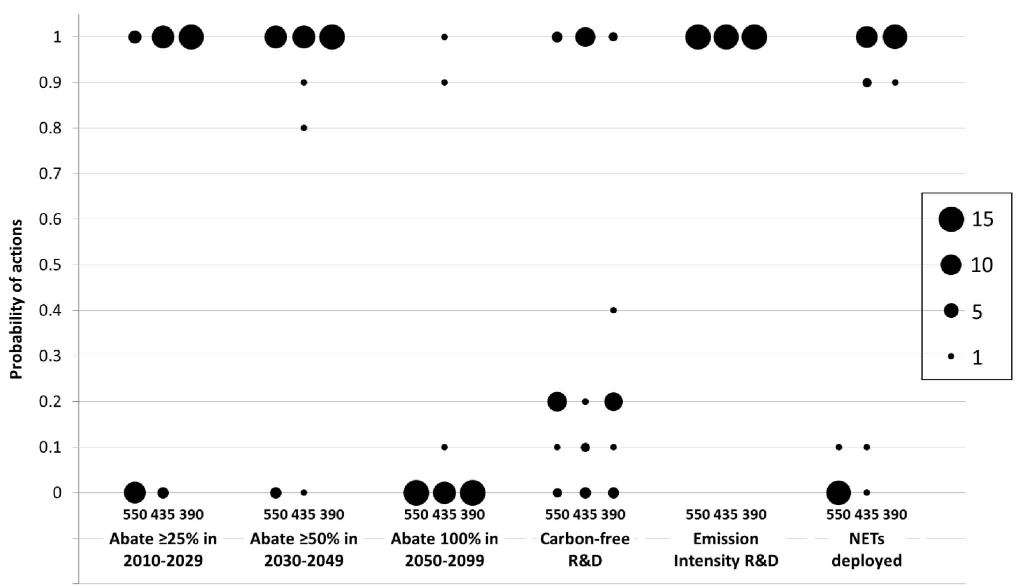
Some policy choices are not sensitive to climate targets or to parameters' values. For example, the optimal portfolio usually abates at least 50% of period 2 BAU emissions and at least 75% of period 3 BAU emissions in nearly all parameterizations (Figures 5.2 and 5.3). Furthermore, public funding for R&D is rarely above half of the maximal level.[10] Unless the CO_2 constraint is a strict threshold or there is no discounting, NETs are almost never used before period 3 or without previous NET R&D. A robust course of action therefore plans for deep abatement from 2030-2100, includes public R&D support that is significant but not a substitute for early abatement, and deploys NETs only after deep abatement and in conjunction with ongoing deep abatement.

The outliers in Figure 5.3 tell their own interesting stories. First, carbon-free R&D is the only category of action that often occurs with a probability strictly between 0 and 1. This happens when NETs are available or when the CO_2 constraint is at its least stringent (550 ppm). Each of these cases requires relatively low future abatement, leading the policymaker to fund emission intensity R&D in period 1. Carbon-free R&D then occurs in period 2 if the emission intensity R&D from period 1 did not have much success and period 1 abatement did not induce much technological change. Second, cases with uncertain NET deployment reveal the interaction of stochastic technology and abatement cost. Increasing the scope for technological change decreases the probability of NET deployment to 0.9 with the 390 ppm constraint and to 0.1 with the 435 ppm constraint because it decreases the expected cost of period 3 abatement. The parameterizations with limited control over technological change and with low-cost abatement also decrease the probability of NET use with the

[10]The main exceptions with public R&D commonly at 75% of the maximal level are: period 2 carbon-free R&D in scenarios with the 435 ppm CO_2 constraint and unavailable NETs, period 2 emission intensity R&D in scenarios with NET options and cheap R&D or cheap abatement, and period 2 NET R&D in scenarios with the 435 or 390 ppm CO_2 constraints.



(a) NETs unavailable



(b) NETs available

Figure 5.3: The probability of undertaking a type of action in each parameterization. For each category of action, the three columns represent the 550 ppm CO_2 constraint (left), the 435 ppm CO_2 constraint (middle), and the 390 ppm CO_2 constraint (right). Each probability is rounded to the nearest multiple of 0.1, and each circle has an area proportional to the percentage of the parameterizations that produce that rounded probability (n=14 without NETs and n=15 with NETs). Probability calculations use, first, the probability of each technology outcome conditional on previous actions and, second, the optimal actions conditional on each set of technology realizations and previous emissions.

435 ppm constraint. Relatedly, the two isolated parameterizations that generally produce 100% abatement in period 3 with the 435 ppm constraint are those with greater scope for technological change and with low-cost abatement. These two parameterizations lead optimal policy to forgo NET deployment in favor of increasing period 3 abatement.

The final insights from the outliers in Figure 5.3 concern the effect of ITC. In the three parameterizations that make ITC stronger, abatement is more successful at producing technological change independently of public R&D policies. This pushes policy in at least three directions. First, near-term abatement could increase as it now provides an additional benefit. Second, near-term R&D could decrease since abatement has become relatively more effective at producing technological change. Third, near-term abatement could decrease while long-term abatement increases so as to take advantage of greater expected technological change. We see the first two effects in our results. In a world with a 390 ppm CO_2 constraint but without NETs, the parameterization with "perfect" ITC is the only one that does not have public funding for carbon-free R&D. Furthermore, variations in the effectiveness of ITC generally also account for the minor variation in the level of public R&D funding. More interestingly, when the CO_2 constraint is at its least stringent so that there is more room to reallocate abatement over time, we see greater near-term abatement in the parameterizations with better ITC. This occurs in order to improve technology for use in later abatement efforts. In sum, we see stronger ITC decreasing R&D funding and increasing near-term abatement, but this only happens under some conditions because ITC is not a dominant factor in our parameterizations.

Finally, the policy environment and emission constraint determine many of the remaining details about the optimal course of action, regardless of the parameter values examined. In a world without NETs, one can almost perfectly predict each period's abatement if one knows the CO_2 constraint and nothing else about the parameterization under consideration. The availability of NETs tends to reduce the importance of the CO_2 constraint for the determination of abatement levels and abatement R&D decisions because NETs can make the more stringent constraints' abatement goals behave more like those needed for less stringent constraints. In a world without NETs, the emission target selected for climate policy almost completely determines immediate abatement and R&D decisions, and in a world with NETs, the emission target determines whether NETs are relevant. Further the possibility of NET use allows the precise level of period 3 abatement (as opposed to the broad categories in Figure 5.3) under the two more stringent CO_2 constraints to be contingent on abatement R&D outcomes and on NET R&D outcomes. For instance, if abatement R&D is not successful while NET R&D is successful, NET deployment can be scaled up and abatement can be scaled down. Because they reduce the probability of undertaking the deepest levels of period 2 and period 3 abatement, available NETs reduce the incentive to invest in carbon-free R&D and increase the incentive to invest in emission intensity R&D. NETs and emission intensity R&D thus act as complements, both substituting for carbon-free R&D and for abatement. Carbon-free R&D is driven by anticipation of deep abatement in the future, and abatement is driven by the cumulative emission constraint. However, NETs effectively truncate abatement cost: they substitute for all abatement beyond their marginal cost, and they therefore increase the value of R&D into emission intensity technology that more strongly affects cost when abatement is lower.

5.4 Discussion: Policy implications

The emission paths (Figure 5.2) and the probability of future deep abatement (Figure 5.3) show that cost-minimizing climate policy portfolios emphasize abatement of 50-100% by 2050 in nearly all parameterizations and under almost any combination of CO_2 targets and available policy options. These levels of medium-term abatement are consistent with the most ambitious goals announced by major emitters (e.g., UNFCCC, 2011). The optimal level of near-term abatement is sensitive to CO_2 targets and to judgments about NETs' cost, risk, and availability, but it is not sensitive to the availability of policies that aim to directly spur clean energy R&D. Because the translation of emissions into temperature is uncertain, announced 2°C temperature limits can only be met with some probability (e.g., Meinshausen et al., 2009). If policymakers accept that the target may be met with less than a 50% chance, then our middle emission constraint might be adopted. In this case, announced 2°C temperature limits imply that emissions over the next half-century should be 50% lower than the BAU path. If policymakers require a greater than 50% chance of meeting the 2°C temperature limit, then our most stringent emission constraint is the more relevant one. With this most stringent constraint, either maximal abatement effort must begin immediately or the policymaker limits nearer-term emission reductions to around 50% while planning for prodigious deployment of NETs later in the century (compare van Vuuren and Riahi, 2011).

While the availability of technology policies generally does not affect abatement paths, these policies can greatly reduce the cost of the optimal policy portfolio (Figure 5.1). Technology policies should emphasize carbon-free technologies if large-scale NET deployment is not viable even as science and policy call for stricter emission constraints; technology policies should emphasize emission intensity technologies if NETs are expected to play a large role later this century. Technology policies are not guaranteed to succeed, but their payoffs are asymmetric: failure leaves future abatement cost unchanged while success lowers it (Bosetti and Tavoni, 2009). Importantly, the effect of technology policies and of NETs on the cost of the policy portfolio has additional significance in a benefit-cost setting where a lower policy cost for a given climate outcome can justify reducing endogenous cumulative emissions.

We only consider worlds with and without NETs, but policymakers in a world without NETs might be able to purchase them. Two factors increase the value of purchasing a NET option. First, if emission constraints are stringent, then we have seen that NETs significantly reduce the cost of the optimal policy portfolio. Second, if the policymaker expects to learn about climate change over time, then NETs' ability to make emissions at least partially reversible confers greater ability to take advantage of future learning (compare Pindyck, 2002; Fisher and Narain, 2003). Different policy instruments provide different incentives for NET development: a cap-and-trade program only values NETs as offsets and provides no incentive for net negative emissions over a trading period, while a carbon tax can incentivize net negative emissions if deployed NETs receive tax credits or carbon payments (i.e., if a linear tax is linear over the whole range of emissions rather than only over positive emissions). The importance of incentives for NET development depends in part on the value of obtaining a NET option.

Three types of research could improve our model's applicability. First, near-term interdisciplinary research into the possible costs, scale, and land use implications of NETs could not only improve the current model but could enable future policy decisions to respond to the new information about NETs (e.g., van Vuuren et al., 2009; Luckow et al., 2010; van Vuuren et al., 2010a). In fact, R&D to reduce NETs' cost from the baseline estimate almost always precedes deployment of NETs in the current model, though it does not appear to be necessary for such deployment. Because NET deployment is valuable for its ability to substitute for abatement, our model's results are necessarily sensitive to NETs' capacity and to the relative cost of abatement and of NET deployment. Second, different functions for probabilistically connecting R&D support and abatement policies to technological outcomes could provide more realistic representations of technological change. These connections could be developed by expert elicitation (e.g., Baker et al., 2009) or by extrapolation from past experience (e.g., Nemet, 2006). However, any such function will remain subject to substantial structural uncertainty when applied to the attempt to shape future energy technology. This observation leads to the third important research path: the portfolio selection model might produce stronger and more detailed policy implications if, beyond its current consideration of parametric uncertainty, it also accounted for structural uncertainty about functional forms and probability distributions. This kind of sensitivity analysis would require either a simpler model or a larger cluster of computers to run the current model, but especially if evaluated with algorithms for supervised learning (Hastie et al., 2009), it could provide a more complete depiction of the connection between policy outcomes and possible factors governing abatement cost and technological change.

The climate and the economy are both complex systems whose evolution over century-long timescales is subject to particularly difficult forms of uncertainty. However, because CO_2 accumulates in the atmosphere and investments in energy infrastructure tend to be long-lived, the climate policies adopted over the next decades play a large role in determining how much flexibility remains later in the century to respond to the realized climate and economy. Any climate policy portfolio implicitly places bets on the climatic and economic systems, but some portfolios imply more specific bets than do others and impose greater costs if their bets turn out poorly. We have taken a step towards representing the policy implications of different types of bets and towards determining which policies cohere with the broadest range of bets. We find that deep intermediate- and long-term abatement is robust to the scenarios considered here, but near-term abatement and R&D funding decisions depend on CO_2 goals and on the anticipated availability of NETs. NETs affect optimal abatement paths if the

CO_2 target is near or below present CO_2 concentrations. In that case, these options can greatly reduce the social cost of the policy portfolio, and they shift some near-term funding for abatement and for carbon-free R&D into funding for R&D targeted towards emission intensity technologies and towards reducing the cost of NETs. Future NET deployment can greatly facilitate the achievement of stringent CO_2 targets and can partially compensate for excess emissions over the next years (Obersteiner et al., 2001). However, depending on future large-scale NET use can be a brittle strategy: NETs and long-term CO_2 storage carry their own risks, and future use of NETs may not help with nearer-term climate thresholds and other irreversible changes. The availability of NETs provides a valuable option to partially undo previous emissions, but abatement also gains option value from increasing future flexibility to forgo reliance on NETs if the technology or climate prove problematic in the interim.

5.5 Appendix: Model parameterization

This appendix describes the parameterization of the portfolio selection model. It describes the probability distributions for technological outcomes, the functional representation of induced technological change (ITC), and the cost functions used in the objective function.

The state variables α_t, γ_t, and ϕ_t record the technology outcomes that apply to period t (Table 5.1). These outcomes are each drawn from a three-point probability distribution similar to the one in Baker and Adu-Bonnah (2008). The main differences are that here the distribution is anchored by the previous period's realized outcome and that here the targeted level of technology depends not just on the previous period's R&D funding but also on its abatement policy. Abatement can induce technological change via functions $ITC_\alpha : \mu_t \rightarrow [0, \bar{\alpha}]$ for carbon-free R&D and $ITC_\gamma : \mu_t \rightarrow [0, \bar{\gamma}]$ for emission intensity R&D. ITC may occur through private R&D or through learning-by-doing. The technology target for a given period comes from summing the targets produced by abatement via ITC and by public R&D, provided the total target does not exceed the exogenous maximal level. The three possibilities are that technology does not change, that the technology target is attained, and that the technology target is surpassed to yield the best possible outcome:[11]

$$Pr[\alpha_t = \alpha_{t-1}] = p_\alpha(1 - \min[\hat{\alpha}_{t-1} + ITC_\alpha(\mu_{t-1}), \bar{\alpha}]) \tag{5.3}$$
$$Pr[\alpha_t = \min(\hat{\alpha}_{t-1} + ITC_\alpha(\mu_{t-1}), \bar{\alpha})] = 1 - p_\alpha \tag{5.4}$$
$$Pr[\alpha_t = \bar{\alpha}] = p_\alpha(\min[\hat{\alpha}_{t-1} + ITC_\alpha(\mu_{t-1}), \bar{\alpha}]) \tag{5.5}$$
$$Pr[\gamma_t = \gamma_{t-1}] = p_\gamma(1 - \min[\hat{\gamma}_{t-1} + ITC_\gamma(\mu_{t-1}), \bar{\gamma}]) \tag{5.6}$$
$$Pr[\gamma_t = \min(\hat{\gamma}_{t-1} + ITC_\gamma(\mu_{t-1}), \bar{\gamma})] = 1 - p_\gamma \tag{5.7}$$
$$Pr[\gamma_t = \bar{\gamma}] = p_\gamma(\min[\hat{\gamma}_{t-1} + ITC_\gamma(\mu_{t-1}), \bar{\gamma}]) \tag{5.8}$$
$$Pr[\phi_t = \phi_{t-1}] = p_\phi(1 - \hat{\phi}_{t-1}) \tag{5.9}$$
$$Pr[\phi_t = \hat{\phi}_{t-1}] = 1 - p_\phi \tag{5.10}$$
$$Pr[\phi_t = \bar{\phi}] = p_\phi \hat{\phi}_{t-1} \tag{5.11}$$

The ITC functions allow us to see how beliefs about the effectiveness of abatement at producing each type of technological change affect optimal policy. Unfortunately, the relationship between ITC and public R&D cannot be specified using empirical results (Pizer and Popp, 2008). Instead, we translate the fraction of emissions abated into the equivalent of some fraction of maximal R&D funding. First, 0% abatement does not affect the R&D targets. Second, we require "perfect" ITC to translate a given percentage abatement into R&D targets that are the same percentage of their maximal levels. This implies that $\mu = ITC_\alpha(\mu)/\bar{\alpha} = ITC_\gamma(\mu)/\bar{\gamma}$ under perfect ITC. A parameter ν controls the effectiveness of ITC and proxies for the severity of innovation market failures. If $\nu = 0$, then ITC for that technology is "perfect" in the sense that a percentage of full abatement produces an equivalent percentage of the maximal technology target. If $\nu > 0$, then ITC for that technology is "imperfect" in the sense that a percentage of full abatement translates into a smaller percentage of the maximal technology target:

$$ITC_\alpha(\mu_t) = \max(0, (\mu_t - \nu_\alpha)\bar{\alpha}) \tag{5.12}$$
$$ITC_\gamma(\mu_t) = \max(0, (\mu_t - \nu_\gamma)\bar{\gamma}) \tag{5.13}$$

When ν_α and ν_γ are positive, abatement may not produce any ITC unless it reaches a sufficiently high level. This representation enables us to vary the effectiveness of ITC across scenarios and also to make ITC more effective for emission intensity technologies than for carbon-free technologies. Under the assumption that emission intensity technologies represent incremental changes that are more responsive to carbon price signals, the base case parameterization assumes that ITC is stronger for emission intensity technologies than for carbon-free technologies.

It remains to define cost functions for abatement, NET deployment, and public R&D targets. First, the cost of abatement depends on the level of abatement and on available

[11]In the case that $\hat{\alpha}_{t-1} + ITC_\alpha(\mu_{t-1}) > \bar{\alpha}$, we have $Pr[\alpha_t = \bar{\alpha}] = (1 - p_\alpha) + p_\alpha\bar{\alpha}$, implying that either $\alpha_t = \bar{\alpha}$ or $\alpha_t = \alpha_{t-1}$. An analogous caveat holds for the probability distribution for γ.

technologies. $c(\mu_t, \alpha_t, \gamma_t)$ is the average cost in the base case of abating fraction μ_t of BAU emissions e_t given R&D outcomes α_t and γ_t:

$$c(\mu_t, \alpha_t, \gamma_t) = \begin{cases} \min\left[\frac{z_t}{\mu_t} d\left(z_t\right), (1-\alpha_t) d(\mu_t)\right] & \text{for base case abatement cost} \\ \min\left[\frac{z_t}{\mu_t} \tilde{d}\left(z_t\right), (1-\alpha_t) \tilde{d}(\mu_t)\right] & \text{for low-cost abatement} \end{cases} \tag{5.14}$$

where $z_t \equiv \max\left[(\mu_t - \gamma_t)/(1-\gamma_t), 0\right]$ as in Baker and Adu-Bonnah (2008). The top expression holds for the base case parameterization and for most others, but the two parameterizations with low-cost abatement use the bottom expression. Both expressions give abatement cost with time t technology as a function of abatement cost with initial technology, but they differ in the function ($d(\cdot)$ or $\tilde{d}(\cdot)$) used to assign the cost with initial technology. In either case, zero abatement costs nothing ($d(0) = \tilde{d}(0) = 0$), and the normalization is $d(1) = 100$. The range of $c(\cdot)$ is [0,100]. The two terms inside the minimization operators give the effect of emission intensity technologies and carbon-free technologies, and the use of the minimization operator assumes that the cheapest type of technology is used at each level of abatement (compare Blyth et al., 2009). Hoogwijk et al. (2008) reported the carbon price yielding aggregate global abatement of 25% to be between \$10/tCO_2 and \$40/tCO_2 and the carbon price yielding aggregate global abatement of 50% to be between \$60/tCO_2 and some level well above \$100/tCO_2. We develop the base case and the low-cost average cost representations by assuming that marginal costs follow a geometric progression at the discretized points and increase linearly between those points.[12] This yields the normalized values:

$$\text{Base case: } d(0.25) = 2.4,\ d(0.50) = 8.4,\ d(0.75) = 28,\ d(1) = 100$$
$$\text{Low-cost: } \tilde{d}(0.25) = 2.4,\ \tilde{d}(0.50) = 6.0,\ \tilde{d}(0.75) = 12,\ \tilde{d}(1) = 27$$

When z_t falls between the above discretization for μ, we define the cost function by assuming average cost is linear between these discretized points. We only model endogenous technological change, so abatement cost does not change unless carbon-free or emission intensity technology changes as described in equations (5.3) through (5.11).

A second type of cost function applies to deployment κ_t of NETs. We represent NETs as having constant marginal cost, which is determined by adjusting the base case average cost of an exogenous level x of period 1 abatement for the outcome ϕ_t of NET R&D:

$$f(\kappa_t, \phi_t) = \kappa_t (1 - \phi_t)\, d(x) \tag{5.15}$$

[12] More specifically, we develop the two marginal cost representations by assuming that: the carbon prices reported in Hoogwijk et al. (2008) represent the marginal cost of abatement; abatement of 25% has a marginal cost of \$20/tCO_2; abatement of 50% makes marginal costs either quintuple (base case) to \$100/tCO_2 or triple (low-cost case) to \$60/tCO_2; higher levels of abatement follow the same geometric progression; and the marginal cost of abating a given fraction of contemporary emissions is unaffected by previous periods' abatement except through modeled technological change.

Converted to non-normalized costs, $x = 0.75$ in a low-cost parameterization corresponds to NETs costing \$115/tCO_2, which is near the low end of recent estimates, and $x = 1$ in the base case parameterization corresponds to NETs costing \$415/tCO_2, which is above many recent estimates (e.g., Rhodes and Keith, 2005; Keith et al., 2006; Uddin and Barreto, 2007; Stolaroff et al., 2008; Keith, 2009; Pielke Jr., 2009).

Finally, a third type of cost function determines how much R&D funding it takes to select a technology target. We assume that the funding that it takes to aim for the chosen public target depends not on the level of the target but on the percentage of the maximal target that it represents. We treat the cost of reaching a percentage of the maximal level of R&D as being an exogenous fraction (specifically: y_g, $y_h y_g$, or y_j) of the base case cost for abating the same percentage of period 1 emissions:

$$g\left(\frac{\hat{\alpha}_t}{\bar{\alpha}}\right) = y_g * d\left(\frac{\hat{\alpha}_t}{\bar{\alpha}}\right) * \frac{\hat{\alpha}_t}{\bar{\alpha}} * e_1 \tag{5.16}$$

$$h\left(\frac{\hat{\gamma}_t}{\bar{\gamma}}\right) = y_h * g\left(\frac{\hat{\gamma}_t}{\bar{\gamma}}\right) \tag{5.17}$$

$$j\left(\frac{\hat{\phi}_t}{\bar{\phi}}\right) = \frac{y_j}{y_g} * g\left(\frac{\hat{\phi}_t}{\bar{\phi}}\right) \tag{5.18}$$

We represent carbon-free R&D costs in terms of average abatement cost because this provides a natural reference point while satisfying the desired property of decreasing returns, and we define the cost of emission intensity R&D as some fraction y_h of the cost of carbon-free R&D. We make abatement cost and R&D cost of similar magnitude because, first, we are looking at the cost of shifting the whole abatement cost curve and, second, we aim to gain more insight than would be obtained by making R&D very cheap or very expensive.

The base case parameters in these functions and probability distributions are chosen to represent values that accord with intuition about, for instance, emission intensity technology being more responsive to abatement than is carbon-free technology. This model's parameterizations are used as demonstrations to aid intuition and to provide a framework for assessing the implications of different beliefs; the results should not be read as either predictive or prescriptive. 14 alternative parameterizations reflect different beliefs about technological change, cost functions, or discounting (Table 5.3). If all parameterizations produce similar results, then we have more confidence that the results are robust to specific values. A more thorough assessment of robustness should also include structural variation in, for instance, the form of the cost functions, the form of the ITC functions converting abatement into R&D targets, and the form of the probability distribution for technological change.

Table 5.3: The 15 parameter scenarios explored with the numerical model. We run each scenario with each possible combination of the three CO_2 constraints, NET availability and unavailability, and emission overshoots allowed and disallowed (Table 5.2).

Scenario	Parameter values	Base case values
Base case	–	–
Low-cost abatement	$\tilde{d}(\cdot)$	$d(\cdot)$
Low-cost R&D	$y_g = y_j = 0.25$	$y_g = y_j = 0.50$
Low-cost emission intensity R&D	$y_h = 0.50$	$y_h = 1$
Low-cost abatement, R&D, and NETs	$\tilde{d}(\cdot), x = 0.75, y_g = y_j = 0.25$	$d(\cdot), x = 1, y_g = y_j = 0.50$
Limited scope for technology	$\bar{\alpha} = \bar{\gamma} = \bar{\phi} = 0.25$	$\bar{\alpha} = \bar{\gamma} = \bar{\phi} = 0.75$
Greater scope for technology	$\bar{\alpha} = \bar{\gamma} = \bar{\phi} = 0.95$	$\bar{\alpha} = \bar{\gamma} = \bar{\phi} = 0.75$
Limited control over technology	$p_\alpha = p_\gamma = p_\phi = 0.75$	$p_\alpha = p_\gamma = p_\phi = 0.25$
High discounting	$\beta = 0.90$	$\beta = 0.95$
No discounting	$\beta = 1$	$\beta = 0.95$
Perfect ITC for both technologies	$\nu_\alpha = \nu_\gamma = 0$	$\nu_\alpha = 0.50, \nu_\gamma = 0.25$
Better ITC for both technologies	$\nu_\alpha = 0.25, \nu_\gamma = 0$	$\nu_\alpha = 0.50, \nu_\gamma = 0.25$
Better ITC for intensity technology	$\nu_\gamma = 0$	$\nu_\gamma = 0.25$
No ITC	$\nu_\alpha = \nu_\gamma = 1$	$\nu_\alpha = 0.50, \nu_\gamma = 0.25$
Low-cost NETs	$x = 0.75$	$x = 1$

Chapter 6

High temperature impacts are a dominant uncertainty in the evaluation of climate targets[1]

The unknown degree to which warming increases damages is more important for the economic evaluation of 21st century climate targets than are temperature uncertainty or abatement cost uncertainty. A quadratic damage function supports a 450 ppm CO_2 target if annual damages from 2.5°C of warming are greater than 0.5% of global output, and it supports a 550 ppm CO_2 target if they are greater than 0.2% of global output. With cubic damages, the 450 ppm and 550 ppm CO_2 targets are supported if 2.5°C of warming costs more than 0.2% and 0.1% of global output, respectively.

Climate policy must confront at least three sources of uncertainty: uncertainty about the warming produced by emission paths, uncertainty about the damages incurred by each degree of warming, and uncertainty about the cost of reducing greenhouse gas emissions. The complexity of the climate and the economy limit our ability to model each system, and anticipated structural changes in each system limit our ability to learn from history. The current climate forcing experiment is without direct analogy in the earth record (Joos and Spahni, 2008; Archer et al., 2009), and climate policy aims to undertake a technology experiment that is, by design, also without direct analogy (Hoffert et al., 2002; Popp, 2010). Scientists fear that we are driving the climate system out of sample, and policy aims to do the same for energy technology. If climate and technology remained within the bounds of experience,

[1]Thanks to Haewon McJeon for providing data.

we could characterize uncertainty using knowledge gleaned from experience, but when either is out of sample, we can only learn so much about the future from the past.

Uncertainty has played many roles in discussion of climate change: as an argument against its relevance (Oreskes and Conway, 2010), as an argument for caution before committing to emission reductions (e.g., Economist, 1997; Bush, 2001), and as an argument for emission reductions as insurance against extreme outcomes (e.g., Holdren, 2009; Economist, 2010). In the academic literature, uncertainty has complex effects on models of climate policy, depending both on the types of irreversibility considered and on the structure of uncertainty (Pindyck, 2007; Weitzman, 2009). Of particular note, integrated assessment models' estimated social cost of carbon are sensitive to several uncertain parameters, including climate sensitivity, the shape of the damage function, the discount rate, and the possibility of climate tipping points (Newbold and Daigneault, 2009; Ackerman et al., 2010; Gerst et al., 2010; Lemoine and Traeger, 2010). Making the modeled policymaker aware of parametric uncertainty and of the potential for learning about it further changes optimal emissions (e.g., Kelly and Kolstad, 1999; Keller et al., 2004; Newbold and Daigneault, 2009). We here consider how key uncertainties interact to affect the net benefits from 450 parts per million (ppm) and 550 ppm targets for the carbon dioxide (CO_2) concentration. By combining different beliefs about temperature change and damages with a set of technology futures, we show that the net benefits of climate targets are primarily sensitive to uncertainty about the damages caused by warming.

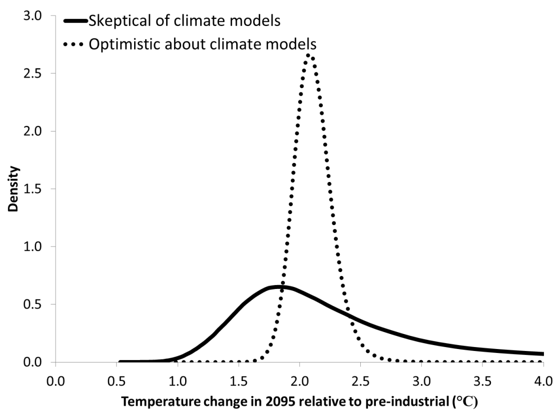
6.1 Three sources of uncertainty

The first source of uncertainty is about the magnitude of warming an emission path causes. This uncertainty can be described by uncertainty about the strength of feedbacks in the earth system (Roe and Baker, 2007). The direct effect of increasing CO_2 is to trap outgoing infrared radiation, requiring the earth's surface to warm in order to restore the balance at the top of the atmosphere between incoming solar energy and outgoing energy. If temperature change were wholly determined by this effect, our uncertainty about its magnitude would be trivial (Pierrehumbert, 2011). However, this change in surface temperature causes other changes in the earth system that themselves affect temperature. For instance, warmer temperatures melt sea ice, replacing white, reflective ice with dark, absorbing ocean. This effect reduces the earth's albedo and further increases surface temperature. In another example, the atmosphere's temperature controls the quantity of water vapor, a potent greenhouse gas (Lacis et al., 2010). Importantly, these and other feedbacks may respond linearly to temperature change but affect temperature nonlinearly (Roe, 2009).

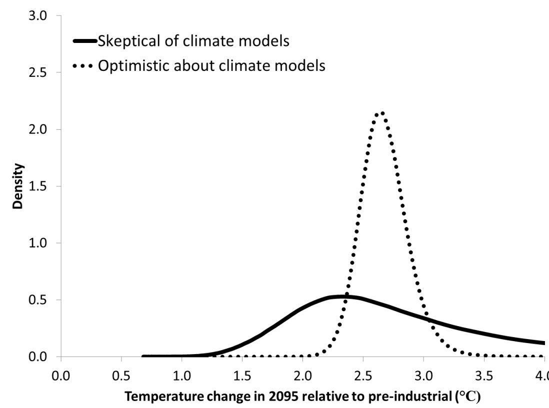
The primary estimates of feedback strength come from climate models, which share knowledge bases, parameterizations, code, and even creators (Tebaldi and Knutti, 2007). Combining models' estimates of feedback strength in a linear feedback framework produces probability distributions like those in Figure 6.1 (Lemoine, 2010a). The dotted lines represent beliefs about temperature change derived from assuming, first, that one model's performance provides no information about other models' performance and, second, that each model includes all relevant feedbacks. This is an optimistic view of models and of our ability to learn from them. In this case of minimized uncertainty, we would be sure that doubling CO_2 produces about 3°C of temperature change. In contrast, the solid lines depict beliefs about temperature change that follow from allowing that models might be biased in common ways and might be incomplete. This skeptical view of models places more weight on extreme temperature outcomes, allowing more than a 5% chance that doubling CO_2 warms the surface by more than 5°C. The results of additional models might now tell us more about the models' common structure than about the real-world processes they aim to represent. When models' errors are correlated, we cannot use them to cross-validate each other.

The second source of uncertainty is the degree to which higher temperatures cause greater impacts. Policy-optimizing integrated assessment models commonly represent warming as reducing economic output by a fraction that increases with a polynomial of temperature (e.g., Nordhaus, 1992, 2008). This polynomial is usually assumed to be quadratic (e.g., Nordhaus, 2008), though some vary the exponent on temperature from 1 (linear) to 3 (cubic) (Dietz et al., 2007). Once the functional form is assumed, its coefficients are estimated using impact assessments for moderate levels of temperature change along with an adjustment for the chance of extreme outcomes (Nordhaus and Boyer, 2000; Hitz and Smith, 2004; Dietz et al., 2007). Without observing how the earth system and the economy respond to temperature increases of various speeds and magnitudes, it is impossible to know how to extrapolate impacts from moderate temperature change to greater temperate change. Many have argued for functional forms that make damages increase faster than does the common quadratic polynomial (e.g., Weitzman, 2009), but the best form or mix of forms will remain a matter of contention for the foreseeable future.

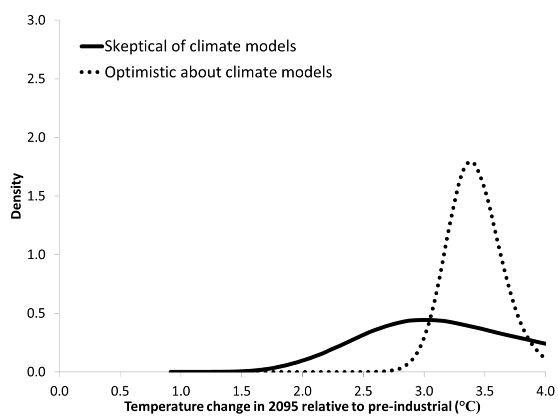
The third source of uncertainty is the cost of reducing emissions to achieve climate targets. Much of this cost will be driven by the cost of replacing electricity generation and transportation infrastructure with lower-carbon substitutes (Barker et al., 2007; Davis et al., 2010). However, the cost of reducing CO_2 emissions depends both on technological possibilities and on the policy path chosen (Hoffert et al., 2002). Because technological change is endogenous, implementing policies today to reduce greenhouse gas emissions should improve the technology available for future policies (Popp et al., 2009), which could strongly affect the cost of future policy. When assessing the cost of reaching long-term climate targets, it is crucial to consider how technology evolves aside from policy and in response to policy (Halsnæs et al., 2007), but this evolution is almost impossible to forecast. Two main methods for describing the possibilities include surveying experts (e.g., Baker et al., 2009) and considering historical experience for similar technologies (e.g., Nemet, 2006). However, neither method is ideal for extrapolating to future technological change. Developing a probability distribution from expert elicitation requires aggregating assessments about breakthroughs which have not yet happened and may not have been conceived. On the other hand, developing a probability distribution for future technological change by extrapolating from historical



(a) 450 ppm in 2095



(b) 550 ppm in 2095



(c) 770 ppm in 2095 (BAU)

Figure 6.1: The distribution of temperature outcomes in 2095 using forcing and transient feedback from GCAM (McJeon et al., in press) and using the "skeptical" and "optimistic" distributions for equilibrium feedback strength (Lemoine, 2010a).

experience either assumes a common structure for technological change across industries or across time within an industry. Yet energy technologies face many challenges defined by chemistry, biology, and physics, and the potential for progress in tackling these challenges is not clearly linked to other industries' successes with their own challenges or to past successes in overcoming energy technologies' previous barriers. Because technological change affects the cost of reducing emissions from each sector, it is more feasible to sample from a range of possible technology outcomes to obtain a range of abatement cost outcomes (e.g., McJeon et al., in press) than it is to take the next step of attaching probabilities to these outcomes.

6.2 Evaluating climate policy under uncertain technology, impacts, and science

When dealing with multiple uncertainties having ambiguous structures interacting via complex systems, it is helpful to consider which types of world are compatible with different actions rather than focusing on determining the optimal action given a set of beliefs (Lempert and Schlesinger, 2000). The latter strategy assesses the optimality of actions under one set of beliefs while the former assesses their brittleness across beliefs. We show how beliefs about the sensitivity of the climate to emissions, the sensitivity of the economy to the climate, and the cost of reducing emissions combine to affect the payoffs from holding year 2100 CO_2 concentrations to 450 ppm and 550 ppm, as opposed to allowing CO_2 concentrations to rise to around 770 ppm in year 2100 under the business-as-usual (or reference) scenario.

There are four main components to this analysis: the distribution for abatement cost, the distribution for equilibrium temperature (i.e., for climate sensitivity), the translation of equilibrium temperature change into temperature change in each period, and the translation of each period's warming into economic damages. Figure 6.2 is a graphical illustration for a simpler static setting. To estimate abatement cost, we turn to a recent combinatorial analysis of technology futures using the Global Change Assessment Model (GCAM) with year 2095 CO_2 limits of 450 and 550 ppm (McJeon et al., in press). This analysis does not give a probability distribution; instead, its output indicates the frequency of abatement cost outcomes among the chosen technological narratives. The present value of abatement cost from 2005-2095 under technology future i for CO_2 target x is $\mu_{x,i}$ (Figure 6.3). The positive skew occurs because when technology breakthroughs partially substitute for each other, obtaining either breakthrough alone may reduce abatement costs by as much as would obtaining the combination. For instance, reducing the cost of nuclear power is less valuable when renewable energy is also much cheaper and carbon can be easily captured from coal-fired power plants.

In order to assess the benefits of each climate target, we translate the radiative forcing in each period for a representative time series of the GCAM results into an effect on the world's economic output. Each period's radiative forcing would produce equilibrium temperature

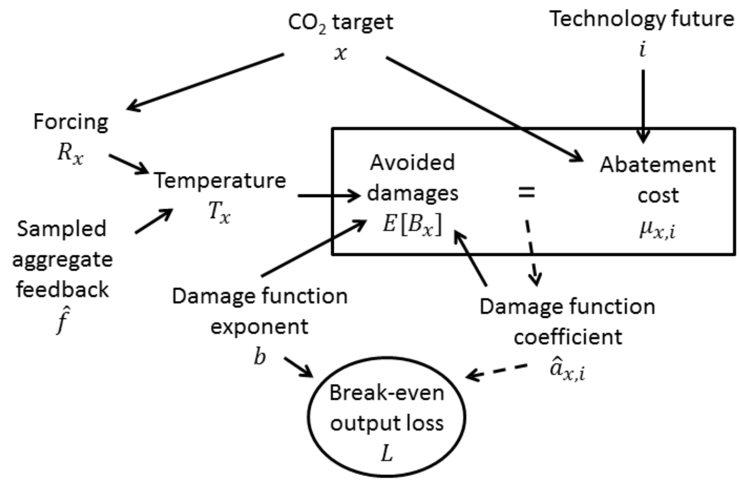


Figure 6.2: A simplified schematic of the calculation of break-even output loss L (in the oval). The two dashed arrows indicate the variables inferred by solving the boxed equation. This figure depicts a static setting, but the actual calculation aggregates benefits over the 21st century and allows for time-varying forcing and transient feedbacks.

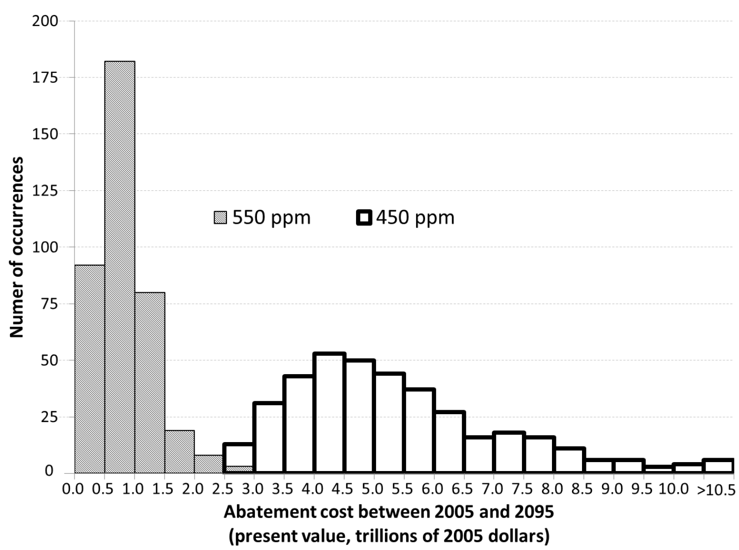


Figure 6.3: The distribution of abatement cost $\mu_{x,i}$ from McJeon et al. (in press) for year 2095 CO_2 targets of 450 ppm and 550 ppm. There are 384 abatement cost outcomes for each target.

change determined by the strength of aggregate feedbacks (Roe, 2009). These aggregate feedbacks have cumulative distribution $G(\hat{f})$, as determined by each of two distributions from Lemoine (2010a): an "optimistic" distribution that assumes climate models are complete and conditionally independent, and a "skeptical" distribution that allows climate models to be incomplete and to share biases. These feedbacks determine equilibrium temperature change, but the ocean heat sink moderates the temperature change actually realized over the 21st century. We represent this ocean heat sink as a time-varying negative feedback $f_{x,t}$ (Baker and Roe, 2009) and calibrate its strength to the temperature outcomes produced by the implementation of the climate model MAGICC in GCAM. The transient negative feedback $f_{x,t}$, a sampled aggregate feedback $\hat{f}$, and time t radiative forcing $R_{x,t}$ along CO_2 path x determine time t temperature change:

$$T_{x,t} = \frac{\lambda R_{x,t}}{1 - \hat{f} - f_{x,t}} \tag{6.1}$$

The parameter λ gives the temperature change per unit of radiative forcing in the reference system upon which feedbacks operate and is approximately 0.315 K $(\mathrm{W\,m^{-2}})^{-1}$ (Soden et al., 2008). Figure 6.1 gives the distribution of temperature in the year 2095 under each CO_2 target.

We translate each temperature change time series into economic impacts for comparison to the cost of abatement. Following the convention in integrated assessment models, we represent damages D as reducing global economic output Y_t by a fraction determined by a polynomial of order b with coefficient a.[2] We take damages to be a function of temperature and the unknown coefficient a for a given polynomial of order b, with no damages occurring if temperature does not change from pre-industrial levels (e.g., Nordhaus, 2008). The economic output remaining after accounting for climate damages is:

$$\hat{Y}_{x,t}(\hat{f}, a) = \frac{Y_t}{1 + D(a, T_{x,t})} = \frac{Y_t}{1 + a[T_{x,t}]^b} \tag{6.2}$$

We explore the implications of using commonly assumed quadratic damages ($b = 2$) and of using a cubic function ($b = 3$) that further increases damages at high temperatures.

The benefit of adopting a climate target is reducing the damage from temperature change, which here means reducing the output lost due to higher temperatures. The present value of this benefit from 2005-2095 is:

$$B_x(\hat{f}, a) = \sum_{t=2005}^{2095} \frac{1}{(1+r)^{t-2005}} \left[Y_t - \hat{Y}_{x,t}(\hat{f}, a) \right] \tag{6.3}$$

[2] In the GCAM runs used in McJeon et al. (in press), output varies between emission paths due to changing use of energy services. However, this is not considered to be a reliable part of the model and should not be combined with the reported abatement cost (personal communication).

The real annual discount rate of $r = 0.05$ matches that used to calculate $\mu_{x,i}$ (McJeon et al., in press). We seek the value of a that equalizes the expected benefit $E[B_x(\hat{f}, a)]$ from achieving target x to each sampled abatement cost $\mu_{x,i}$. At this "break-even" value, an optimizing global policymaker would be indifferent between achieving climate target x or not when abatement cost is known to be at the sampled value $\mu_{x,i}$. The expected benefit is calculated for a given functional form for damages and for a given distribution for aggregate feedbacks. The break-even value $\hat{a}_{x,i}$ is defined implicitly by:

$$\mu_{x,i} = \int B_x(\hat{f}, \hat{a}_{x,i})\ G'(\hat{f})\ d\hat{f} \tag{6.4}$$

If the true value for a is greater than its break-even value $\hat{a}_{x,i}$, then target x provides net benefits in expectation. In order to assess the plausibility of the resulting damage function coefficients, we translate them into the fraction of global output that would be lost annually from temperature change of 2.5°C:

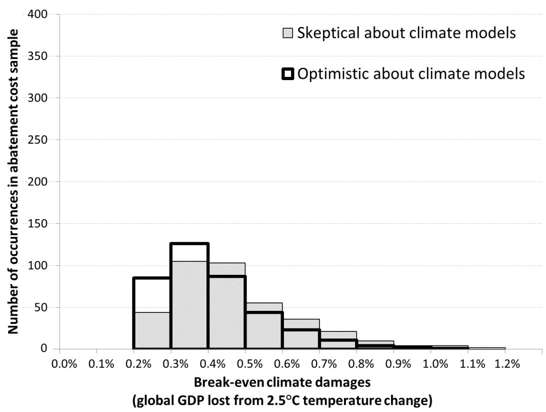
$$L(\hat{a}_{x,i}) = 1 - \frac{1}{1 + \hat{a}_{x,i}[2.5]^b} \tag{6.5}$$

The distributions for L reveal which beliefs about climate damages and technology are consistent with 450 and 550 ppm targets for the year 2100. If the actual output loss from 2.5°C of temperature change is expected to be above the distribution for L, then the CO_2 target looks generally attractive, and if it is expected to be below the distribution for L, then the CO_2 target looks generally unattractive.

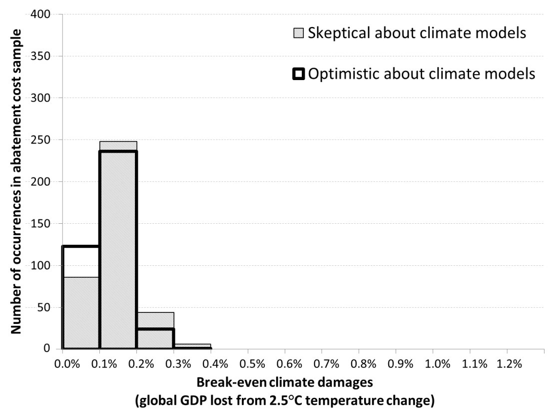
6.3 Results: Break-even output loss

We assess which beliefs about climate damages equalize the expected benefit of 450 ppm and 550 ppm climate targets to the cost of abatement in each technological future. Figure 6.4 plots the distribution of L for each climate target and for quadratic and cubic damage functions. Within a plot, outcomes that have a greater break-even output loss (i.e., outcomes farther to the right) correspond to futures with less progress in abatement technology. Economic studies have generally estimated that about 1-2% of output is lost from 2.5°C of warming, with many of the omitted factors and potential surprises tending to further increase damages (Nordhaus and Boyer, 2000; Tol, 2009).

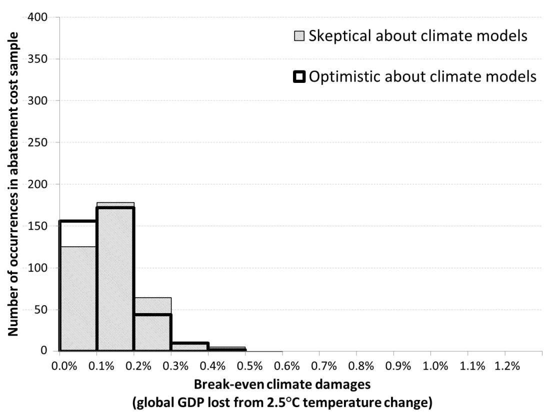
The results for break-even output loss provide three insights. First, the curvature of the damage function is a crucial assumption. When the damage function is quadratic, the break-even output loss for 2.5°C of warming ranges from 0.2% to 1.2% for a 450 ppm target and from 0.03% to 0.5% for a 550 ppm target. Both of these ranges are below economic estimates (Tol, 2009), indicating that stabilizing the CO_2 concentration at these levels in 2095 should provide sufficient benefits to cover its costs. With a cubic damage function that increases the loss from the highest temperature outcomes, the ranges are both lower and narrower:



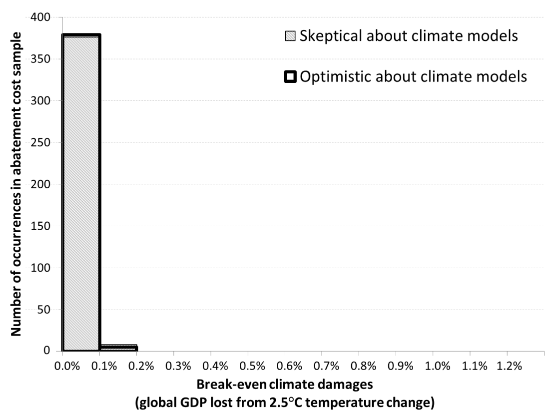
(a) CO_2 target: 450 ppm, Damages: quadratic



(b) CO_2 target: 450 ppm, Damages: cubic



(c) CO_2 target: 550 ppm, Damages: quadratic



(d) CO_2 target: 550 ppm, Damages: cubic

Figure 6.4: The damages L from 2.5°C of warming (as a percentage of global output) that would equalize each abatement cost $\mu_{x,i}$ in Figure 6.3 to the expected benefits $E[B_x(\hat{f}, a)]$ over the next 90 years of adopting CO_2 target x for 2095 ($n = 384$). If the damages from 2.5°C of warming are believed to be less than the break-even level required by a given abatement cost, then the CO_2 target does not pay for itself in expectation over this time horizon, but if damages are thought to be greater, then the CO_2 target does pay for itself in expectation.

with a floor under 0.01% for either target and with a maximum of 0.3% for the 450 ppm target and of 0.1% for the 550 ppm target. More convex damages make the break-even output loss largely insensitive to abatement cost, requiring little damages from moderate warming to justify either CO_2 target. The greater impacts at high temperatures dominate the cost-benefit analysis, and the uncertainty in the assumed shape of the damage function is more important for optimal policy than is uncertainty about abatement technology.

Second, fat-tailed distributions for climate sensitivity have received much attention in the recent literature as an argument for greater emission reductions (e.g., Weitzman, 2009), but while the two distributions give similar results, we find that it is actually the thinner-tailed distribution that supports the analyzed climate policies at less severe damage functions. The reason for this result is twofold. First, the thicker-tailed distribution does place more weight on higher temperature outcomes, but because it allows the group of models to be biased high or low, it also shifts probability mass to lower temperature outcomes. Second, the effect of the ocean's heat capacity in moderating transient warming reduces the influence of extreme feedback outcomes on temperature over the next decades (Baker and Roe, 2009), which decreases the influence of the temperature outcomes in the positive tail of the fatter-tailed "skeptical" distribution. While greater confidence in models' ability to capture all relevant processes in a representative way decreases the probability of extreme temperature outcomes (Figure 6.1), these extreme outcomes matter less over the next decades than does the central tendency of the distribution.

The third insight from Figure 6.4 is that the 550 ppm target is consistent with damage functions that would not be severe enough to support the 450 ppm target. Even if estimated damages support the 550 ppm target, the additional abatement cost required for the 450 ppm target might more than offset the additional reduction in damages. The 450 ppm target is supported for damage functions that are less severe than common estimates, and the 550 ppm target is supported for a still-wider range of damage functions. However, because this analysis is restricted to 21st century effects, much of the additional abatement cost of the more ambitious target is counted while the additional climate benefits largely occur beyond the time horizon of the analysis.

6.4 The interaction of policy options with uncertainty

The results indicate the importance for climate policy evaluations of assumptions about the shape of the damage function. If multiplicative damages increase faster than the square of temperature, then the expected benefits of adopting a 450 ppm CO_2 target are in the same range as the cost of reducing emissions. Damage uncertainty has a greater impact than does abatement cost uncertainty, which has a greater impact than do the assumptions about climate models used to estimate temperature uncertainty. However, three important caveats apply. First, we follow the abatement cost calculations in using a consumption discount rate of 5%. Because of the long timescales involved and the uneven distribution of costs and benefits over time, alternate assumptions would shift the distribution of break-even damages. Second, the feedback framework assumes that the climate system responds smoothly to increasing forcing, and the damage calculations assume that society responds smoothly to increasing temperature, but more abrupt or discontinuous changes are possible (Alley et al., 2003; Overpeck and Cole, 2006; Lenton et al., 2008). Third, this paper only calculates abatement cost and climate benefits out to the year 2095, but both abatement and warming continue beyond that date. While the cost of abatement might decrease over time as technology improves and low-carbon infrastructure is built, the benefits of a lower CO_2 concentration should increase beyond 2095 as damages from higher temperatures continue into the future and as the earth system more fully responds to additional CO_2 forcing (i.e., as the transient feedback becomes less negative). These two factors tend to bias our calculation of break-even output loss to the high side, thereby making climate policy look less attractive than it would appear in a more complete model.

Given the difficult nature of the uncertainty involved, climate policy should be seen as a hedge against a range of outcomes rather than as set using finely calculated probabilities. We have shown that the benefits of climate policy are dominated not by temperature uncertainty but by uncertainty about the degree to which damages increase with temperature. Further, while adopting a climate target exposes the global economy to abatement cost uncertainty, we have also shown that the range of abatement cost outcomes is less important than the impact of plausible assumptions about damages at high temperatures. Indeed, more convex damage functions make policy less sensitive to abatement cost uncertainty. Models of optimal climate policy should therefore focus more on uncertainty about the functional form for damages, and climate policy negotiations should focus both on avoiding worlds that expose society to high temperatures and on mitigating the consequences of high temperature outcomes.

Emission reductions are valuable for protecting society against the possibility that high temperatures cause much higher damages. Ambitious climate policy protects against surprisingly high damages while taking advantage of potential surprises in technological change. Forgoing emission reductions leaves society exposed to high damage outcomes while gambling that stalled technology would make abatement unsupportably expensive. Other policy options also offer means of skewing bets. Funding low-carbon research and development increases the chance that low-cost abatement technology is available when needed. Developing negative emission technologies gives the world the ability to reduce CO_2 concentrations, helping to mitigate high damage outcomes by getting rid of historical emissions. And research into geoengineering could provide a means of stalling the most extreme damage scenarios until negative emission technologies can return CO_2 concentrations to safer levels. Because each policy option stacks the deck in a different way, combining them in a portfolio should further decrease exposure to bad climatic outcomes while increasing exposure to positive technological outcomes.

Bibliography

Abrego, L. and C. Perroni (2002). "Investment subsidies and time-consistent environmental policy." Oxford Economic Papers-New Series 54(4): 617–635. doi:10.1093/oep/54.4.617.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (in press). "The environment and directed technical change." American Economic Review .

Ackerman, F., S. DeCanio, R. Howarth, and K. Sheeran (2009). "Limitations of integrated assessment models of climate change." Climatic Change 95(3-4): 297–315. doi:10.1007/s10584-009-9570-x.

Ackerman, F., E. A. Stanton, and R. Bueno (2010). "Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE." Ecological Economics 69(8): 1657–1665. doi:10.1016/j.ecolecon.2010.03.013.

Allen, M. R., D. J. Frame, C. Huntingford, C. D. Jones, J. A. Lowe, M. Meinshausen, and N. Meinshausen (2009). "Warming caused by cumulative carbon emissions towards the trillionth tonne." Nature 458(7242): 1163–1166. doi:10.1038/nature08019.

Alley, R. B., J. Marotzke, W. Nordaus, J. Overpeck, D. Peteet, R. P. Jr., R. Pierrehumbert, P. Rhines, T. Stocker, L. Talley, J. Wallace, A. Isern, J. Dandelski, C. Elfring, M. Gopnik, M. Kelly, J. Bachim, and A. Carlisle (2002). Abrupt Climate Change: Inevitable Surprises. Washington, DC: National Research Council, National Academy Press.

Alley, R. B., J. Marotzke, W. D. Nordhaus, J. T. Overpeck, D. M. Peteet, R. A. Pielke, R. T. Pierrehumbert, P. B. Rhines, T. F. Stocker, L. D. Talley, and J. M. Wallace (2003). "Abrupt climate change." Science 299(5615): 2005–2010. doi:10.1126/science.1081056.

Amir, R., M. Germain, and V. van Steenberghe (2008). "On the impact of innovation on the marginal abatement cost curve." Journal of Public Economic Theory 10(6): 985–1010. doi:10.1111/j.1467-9779.2008.00393.x.

Annan, J. D. and J. C. Hargreaves (2006). "Using multiple observationally-based constraints to estimate climate sensitivity." Geophysical Research Letters 33: L06704. doi:10.1029/2005GL025259.

Annan, J. D. and J. C. Hargreaves (2011). "On the generation and interpretation of probabilistic estimates of climate sensitivity." Climatic Change 104(3-4): 423–436. doi: 10.1007/s10584-009-9715-y.

Archer, D. (2007). "Methane hydrate stability and anthropogenic climate change." Biogeosciences 4(4): 521–544.

Archer, D., M. Eby, V. Brovkin, A. Ridgwell, L. Cao, U. Mikolajewicz, K. Caldeira, K. Matsumoto, G. Munhoven, A. Montenegro, and K. Tokos (2009). "Atmospheric lifetime of fossil fuel carbon dioxide." Annual Review of Earth and Planetary Sciences 37(1): 117–134. doi:10.1146/annurev.earth.031208.100206.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). "Coherent measures of risk." Mathematical Finance 9(3): 203–228. doi:10.1111/1467-9965.00068.

Azar, C. and K. Lindgren (2003). "Catastrophic events and stochastic cost-benefit analysis of climate change." Climatic Change 56(3): 245–255. doi:10.1023/A:1021743622080.

Azar, C., K. Lindgren, E. Larson, and K. Möllersten (2006). "Carbon capture and storage from fossil fuels and biomass—Costs and potential role in stabilizing the atmosphere." Climatic Change 74(1): 47–79. doi:10.1007/s10584-005-3484-7.

Azar, C., K. Lindgren, M. Obersteiner, K. Riahi, D. van Vuuren, K. den Elzen, K. Möllersten, and E. Larson (2010). "The feasibility of low CO2 concentration targets and the role of bio-energy with carbon capture and storage (BECCS)." Climatic Change 100(1): 195–202. doi:10.1007/s10584-010-9832-7.

Baker, E. and K. Adu-Bonnah (2008). "Investment in risky R&D programs in the face of climate uncertainty." Energy Economics 30(2): 465–486. doi:10.1016/j.eneco.2006.10.003.

Baker, E., H. Chon, and J. Keisler (2009). "Advanced solar R&D: Combining economic analysis with expert elicitations to inform climate policy." Energy Economics 31(Supplement 1): S37–S49. doi:10.1016/j.eneco.2007.10.008.

Baker, E. and E. Shittu (2008). "Uncertainty and endogenous technical change in climate policy models." Energy Economics 30(6): 2817–2828. doi:10.1016/j.eneco.2007.10.001.

Baker, M. B. and G. H. Roe (2009). "The shape of things to come: Why is climate change so predictable?" Journal of Climate 22(17): 4574–4589. doi:10.1175/2009JCLI2647.1.

Bansal, R., D. Kiku, and A. Yaron (2010). "Long run risks, the macroeconomy, and asset prices." American Economic Review 100(2): 542–546. doi:10.1257/aer.100.2.542.

Bansal, R. and A. Yaron (2004). "Risks for the long run: A potential resolution of asset pricing puzzles." The Journal of Finance 59(4): 1481–1509. doi:10.1111/j.1540-6261.2004.00670.x.

Baranzini, A., M. Chesney, and J. Morisset (2003). "The impact of possible climate catastrophes on global warming policy." Energy Policy 31(8): 691–701. doi:10.1016/S0301-4215(02)00101-5.

Barker, T., I. Bashmakov, A. Alharthi, M. Amann, L. Cifuentes, J. Drexhage, M. Duan, O. Edenhofer, B. Flannery, M. Grubb, M. Hoogwijk, F. I. Ibitoye, C. J. Jepma, W. A. Pizer, and K. Yamaji (2007). "Mitigation from a cross-sectoral perspective." In B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, and L. A. Meyer, eds., "Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Benson, S., P. Cook, J. Anderson, S. Bachu, H. B. Nimir, B. Basu, J. Bradshaw, G. Deguchi, J. Gale, G. von Goerne, W. Heidug, S. Holloway, R. Kamal, D. Keith, P. Lloyd, P. Rocha, B. Senior, J. Thomson, T. Torp, T. Wildenborg, M. Wilson, F. Zarlenga, and D. Zhou (2005). "Underground geological storage." In B. Metz, O. Davidson, H. de Coninck, M. Loos, and L. Meyer, eds., "IPCC Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change," pp. 195–276. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Berger, A. (1992). "Astronomical theory of Paleoclimates and the last glacial-interglacial cycle." Quaternary Science Reviews 11(5): 571–581. doi:10.1016/0277-3791(92)90014-Y.

Berger, A. L. (1978). "Long-term variations of daily insolation and Quaternary climatic changes." Journal of the Atmospheric Sciences 35(12): 2362–2367. doi:10.1175/1520-0469(1978)035⟨2362:LTVODI⟩2.0.CO;2.

Berliner, L. M. and Y. Kim (2008). "Bayesian design and analysis for superensemble-based climate forecasting." Journal of Climate 21(9): 1891–1910. doi:10.1175/2007JCLI1619.1.

Bernardo, J. M. and A. F. M. Smith (1994). Bayesian Theory. Chichester and New York: Wiley.

Biglaiser, G., J. K. Horowitz, and J. Quiggin (1995). "Dynamic pollution regulation." Journal of Regulatory Economics 8(1): 33–44. doi:10.1007/BF01066598.

Blackstock, J. J. and J. C. S. Long (2010). "The politics of geoengineering." Science 327(5965): 527. doi:10.1126/science.1183877.

Blyth, W., R. Bradley, D. Bunn, C. Clarke, T. Wilson, and M. Yang (2007). "Investment risks under uncertain climate change policy." Energy Policy 35(11): 5766–5773. doi: 10.1016/j.enpol.2007.05.030.

Blyth, W., D. Bunn, J. Kettunen, and T. Wilson (2009). "Policy interactions, risk and price formation in carbon markets." Energy Policy 37(12): 5192–5207. doi:10.1016/j.enpol.2009.07.042.

Bony, S., R. Colman, V. M. Kattsov, R. P. Allan, C. S. Bretherton, J. Dufresne, A. Hall, S. Hallegatte, M. M. Holland, W. Ingram, D. A. Randall, B. J. Soden, G. Tselioudis, and M. J. Webb (2006). "How well do we understand and evaluate climate change feedback processes?" Journal of Climate 19(15): 3445–3482. doi:10.1175/JCLI3819.1.

Borsuk, M. and L. Tomassini (2005). "Uncertainty, imprecision, and the precautionary principle in climate change assessment." Water Science & Technology 52(6): 213–225.

Bosetti, V., C. Carraro, E. Massetti, and M. Tavoni (2008). "International energy R&D spillovers and the economics of greenhouse gas atmospheric stabilization." Energy Economics 30(6): 2912–2929. doi:10.1016/j.eneco.2008.04.008.

Bosetti, V. and M. Tavoni (2009). "Uncertain R&D, backstop technology and GHGs stabilization." Energy Economics 31(Supplement 1): S18–S26. doi:10.1016/j.eneco.2008.03.002.

Boykoff, M. T., D. Frame, and S. Randalls (2010). "Discursive stability meets climate instability: A critical exploration of the concept of 'climate stabilization' in contemporary climate policy." Global Environmental Change 20(1): 53–64. doi:10.1016/j.gloenvcha.2009.09.003.

Brozović, N. and W. Schlenker (2011). "Optimal management of an ecosystem with an unknown threshold." Ecological Economics 70(4): 627–640. doi:10.1016/j.ecolecon.2010.10.001.

Bush, G. W. (2001). "President Bush discusses global climate change."

Cadule, P., L. Bopp, and P. Friedlingstein (2009). "A revised estimate of the processes contributing to global warming due to climate-carbon feedback." Geophysical Research Letters 36: L14705. doi:10.1029/2009GL038681.

Cheng, H., R. L. Edwards, W. S. Broecker, G. H. Denton, X. Kong, Y. Wang, R. Zhang, and X. Wang (2009). "Ice Age terminations." Science 326(5950): 248–252. doi:10.1126/science.1177840.

Clarke, L., J. Edmonds, V. Krey, R. Richels, S. Rose, and M. Tavoni (2009). "International climate policy architectures: Overview of the EMF 22 International Scenarios." Energy Economics 31(Supplement 2): S64–S81. doi:10.1016/j.eneco.2009.10.013.

Clarke, L., J. Weyant, and A. Birky (2006). "On the sources of technological change: Assessing the evidence." Energy Economics 28(5-6): 579–595. doi:10.1016/j.eneco.2006.05.004.

Clarke, L., J. Weyant, and J. Edmonds (2008). "On the sources of technological change: What do the models assume?" Energy Economics 30(2): 409–424. doi:10.1016/j.eneco.2006.05.023.

Colman, R. (2003). "A comparison of climate feedbacks in general circulation models." Climate Dynamics 20(7): 865–873. doi:10.1007/s00382-003-0310-z.

Colman, R. and B. McAvaney (2009). "Climate feedbacks under a very broad range of forcing." Geophysical Research Letters 36: L01702. doi:10.1029/2008GL036268.

Colman, R. A., S. B. Power, and B. J. McAvaney (1997). "Non-linear climate feedback analysis in an atmospheric general circulation model." Climate Dynamics 13(10): 717–731. doi:10.1007/s003820050193.

Cox, P. and C. Jones (2008). "Illuminating the modern dance of climate and CO2." Science 321(5896): 1642–1644. doi:10.1126/science.1158907.

Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell (2000). "Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model." Nature 408(6809): 184–187. doi:10.1038/35041539.

Creyts, J., A. Derkach, S. Nyquist, K. Ostrowski, and J. Stephenson (2007). *Reducing U.S. greenhouse gas emissions: How much at what cost?* , McKinsey & Company.

Crost, B. and C. P. Traeger (2010). *Risk and aversion in the integrated assessment of climate change*. CUDARE Working Paper 1104, University of California, Berkeley.

Damen, K., A. Faaij, and W. Turkenburg (2006). "Health, safety and environmental risks of underground CO2 storage — Overview of mechanisms and current knowledge." Climatic Change 74(1): 289–318. doi:10.1007/s10584-005-0425-9.

Dasgupta, P., P. Hammond, and E. Maskin (1980). "On imperfect information and optimal pollution control." The Review of Economic Studies 47(5): 857–860.

Davis, S. J., K. Caldeira, and H. D. Matthews (2010). "Future CO2 emissions and climate change from existing energy infrastructure." Science 329(5997): 1330 –1333. doi:10.1126/science.1188566.

de F. Forster, P. M. and M. Collins (2004). "Quantifying the water vapour feedback associated with post-Pinatubo global cooling." Climate Dynamics 23(2): 207–214. doi:10.1007/s00382-004-0431-z.

Dehghani, M. H. (2011). *Policy uncertainty and technology adoption.* Working paper, University of Texas at Austin.

Delaygue, G. and E. Bard (2010). "An Antarctic view of Beryllium-10 and solar activity for the past millennium." Climate Dynamics doi:10.1007/s00382-010-0795-1.

Denman, K., G. Brasseur, A. Chidthaisong, P. Ciais, P. M. Cox, R. E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandran, P. da Silva Dias, S. Wofsy, and X. Zhang (2007). "Couplings between changes in the climate system and biogeochemistry." In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, eds., "Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," pp. 500–587. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Dessler, A. E., Z. Zhang, and P. Yang (2008). "Water-vapor climate feedback inferred from climate fluctuations, 2003-2008." Geophysical Research Letters 35: L20704. doi:10.1029/2008GL035333.

Dietz, S., C. Hope, and N. Patmore (2007). "Some economics of 'dangerous' climate change: Reflections on the Stern Review." Global Environmental Change 17(3-4): 311–325. doi:10.1016/j.gloenvcha.2007.05.008.

Dongen, S. V. (2006). "Prior specification in Bayesian statistics: Three cautionary tales." Journal of Theoretical Biology 242(1): 90–100. doi:10.1016/j.jtbi.2006.02.002.

Dorrepaal, E., S. Toet, R. S. P. van Logtestijn, E. Swart, M. J. van de Weg, T. V. Callaghan, and R. Aerts (2009). "Carbon respiration from subsurface peat accelerated by climate warming in the subarctic." Nature 460(7255): 616–619. doi:10.1038/nature08216.

Economist (1997). "For Kyoto, a modest proposal. Leader." The Economist .

Economist (2010). "Spin, science and climate change. Leader." The Economist .

Edenhofer, O., B. Knopf, T. Barker, L. Baumstark, E. Bellevrat, B. Chateau, P. Criqui, M. Isaac, A. Kitous, S. Kypreos, M. Leimbach, K. Lessmann, B. Magné, Şerban Scrieciu, H. Turton, and D. P. van Vuuren (2010). "The economics of low stabilization: Model comparison of mitigation strategies and costs." Energy Journal 31: 11–48.

Eglin, T., P. Ciais, S. L. Piao, P. Barre, V. Bellassen, P. Cadule, C. Chenu, T. Gasser, C. Koven, M. Reichstein, and P. Smith (2010). "Historical and future perspectives of global soil carbon response to climate and land-use changes." Tellus B 62(5): 700–718. doi:10.1111/j.1600-0889.2010.00499.x.

Ellsberg, D. (1961). “Risk, ambiguity, and the Savage axioms.” The Quarterly Journal of Economics 75(4): 643–669.

Epstein, L. G. and S. E. Zin (1989). “Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework.” Econometrica 57(4): 937–969.

Fankhauser, S. and C. Hepburn (2010). “Designing carbon markets. Part I: Carbon markets in time.” Energy Policy 38(8): 4363–4370. doi:10.1016/j.enpol.2010.03.064.

Field, C. B., D. B. Lobell, H. A. Peters, and N. R. Chiariello (2007). “Feedbacks of terrestrial ecosystems to climate change.” Annual Review of Environment and Resources 32(1): 1–29. doi:10.1146/annurev.energy.32.053006.141119.

Fischer, C. and R. G. Newell (2008). “Environmental and technology policies for climate mitigation.” Journal of Environmental Economics and Management 55(2): 142–162. doi:10.1016/j.jeem.2007.11.001.

Fischer, C., I. W. H. Parry, and W. A. Pizer (2003). “Instrument choice for environmental protection when technological innovation is endogenous.” Journal of Environmental Economics and Management 45(3): 523–545. doi:10.1016/S0095-0696(03)00002-0.

Fisher, A. C. and U. Narain (2003). “Global warming, endogenous risk, and irreversibility.” Environmental and Resource Economics 25(4): 395–416. doi:10.1023/A:1025056530035.

Fisher, B. S., N. Nakićenović, K. Alfsen, J. C. Morlot, F. de la Chesnaye, J. Hourcade, K. Jiang, M. Kainuma, E. L. Rovere, A. Matysek, A. Rana, K. Riahi, R. Richels, S. Rose, D. van Vuuren, and R. Warren (2007). “Issues related to mitigation in the long term context.” In B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, and L. A. Meyer, eds., “Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,” Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Forster, P., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D. Fahey, J. Haywood, J. Lean, D. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz, and R. van Dorland (2007). “Changes in atmospheric constituents and in radiative forcing.” In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, eds., “Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,” pp. 129–234. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

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.” Geophysical Research Letters 32: L09702. doi:10.1029/2004GL022241.

Frank, D. C., J. Esper, C. C. Raible, U. Buntgen, V. Trouet, B. Stocker, and F. Joos (2010). "Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate." Nature 463(7280): 527–530. doi:10.1038/nature08769.

Friedlingstein, P., P. Cox, R. Betts, L. Bopp, W. von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, G. Bala, J. John, C. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H. D. Matthews, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K. Schnitzler, R. Schnur, K. Strassmann, A. J. Weaver, C. Yoshikawa, and N. Zeng (2006). "Climate-carbon cycle feedback analysis: Results from the C$ˆ4$MIP model intercomparison." Journal of Climate 19(14): 3337–3353. doi:10.1175/JCLI3800.1.

Friedlingstein, P., J. Dufresne, P. M. Cox, and P. Rayner (2003). "How positive is the feedback between climate change and the carbon cycle?" Tellus B 55(2): 692–700. doi: 10.1034/j.1600-0889.2003.01461.x.

Gamerman, D. and H. F. Lopes (2006). Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Boca Raton: Chapman & Hall/CRC, 2nd edition.

Gelman, A. (2006). "Prior distributions for variance parameters in hierarchical models (Comment on article by Browne and Draper)." Bayesian Analysis 1(3): 515–534. doi: 10.1214/06-BA117A.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). Bayesian Data Analysis. Boca Raton, Fla: Chapman & Hall/CRC, 2nd edition.

Gerlagh, R., S. Kverndokk, and K. Rosendahl (2009). "Optimal timing of climate change policy: Interaction between carbon taxes and innovation externalities." Environmental and Resource Economics 43(3): 369–390. doi:10.1007/s10640-009-9271-y.

Gersbach, H. and A. Glazer (1999). "Markets and regulatory hold-up problems." Journal of Environmental Economics and Management 37(2): 151–164. doi:10.1006/jeem.1998.1063.

Gerst, M. D., R. B. Howarth, and M. E. Borsuk (2010). "Accounting for the risk of extreme outcomes in an integrated assessment of climate change." Energy Policy 38(8): 4540–4548. doi:10.1016/j.enpol.2010.04.008.

Gjerde, J., S. Grepperud, and S. Kverndokk (1999). "Optimal climate policy under the possibility of a catastrophe." Resource and Energy Economics 21(3-4): 289–317. doi: 10.1016/S0928-7655(99)00006-8.

Goulder, L. H. and K. Mathai (2000). "Optimal CO2 abatement in the presence of induced technological change." Journal of Environmental Economics and Management 39(1): 1–38. doi:10.1006/jeem.1999.1089.

Greaker, M. and L. Pade (2009). "Optimal carbon dioxide abatement and technological change: should emission taxes start high in order to spur R&D?" Climatic Change 96(3): 335–355. doi:10.1007/s10584-009-9643-x.

Greenstone, M., E. Kopits, and A. Wolverton (2011). "Estimating the social cost of carbon for use in U.S. federal rulemakings: a summary and interpretation." National Bureau of Economic Research Working Paper Series No. 16913.

Gregory, J. M., C. D. Jones, P. Cadule, and P. Friedlingstein (2009). "Quantifying carbon cycle feedbacks." Journal of Climate 22(19): 5232–5250. doi:10.1175/2009JCLI2949.1.

Gruber, N., P. Friedlingstein, C. B. Field, R. Valentini, M. Heimann, J. E. Richey, P. R. Lankao, E. Schulze, and C. A. Chen (2004). "The vulnerability of the carbon cycle in the 21st century: An assessment of carbon-climate-human interactions." In C. Field and M. Raupach, eds., "The Global Carbon Cycle: Integrating Humans, Climate and the Natural World," pp. 45–76. Washington, D.C.: Island Press.

Guillerminet, M. and R. S. J. Tol (2008). "Decision making under catastrophic risk and learning: the case of the possible collapse of the West Antarctic Ice Sheet." Climatic Change 91(1): 193–209. doi:10.1007/s10584-008-9447-4.

Hall, D. C. and R. J. Behl (2006). "Integrating economic analysis and the science of climate instability." Ecological Economics 57(3): 442–465. doi:10.1016/j.ecolecon.2005.05.001.

Hall, J. (2007). "Probabilistic climate scenarios may misrepresent uncertainty and lead to bad adaptation decisions." Hydrological Processes 21(8): 1127–1129. doi:10.1002/hyp.6573.

Hallegatte, S., A. Lahellec, and J. Grandpeix (2006). "An elicitation of the dynamic nature of water vapor feedback in climate change using a 1D model." Journal of the Atmospheric Sciences 63(7): 1878–1894. doi:10.1175/JAS3725.1.

Halsnæs, K., P. Shukla, D. Ahuja, G. Akuma, R. Beale, J. Edmonds, C. Gollier, A. Grübler, M. H. Duong, A. Markandya, M. McFarland, E. Nikitina, T. Sugiyama, A. Villavicencio, and J. Zou (2007). "Framing issues." In B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, and L. A. Meyer, eds., "Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Hanemann, M., S. H. Schneider, A. Rosencranz, M. Mastrandrea, and K. Kuntz-Duriseti (2010). "What is the economic cost of climate change?" In "Climate Change Science and Policy," pp. 185–193. Washington, D.C.: Island Press.

Hannart, A., J. Dufresne, and P. Naveau (2009). "Why climate sensitivity may not be so unpredictable." Geophysical Research Letters 36: L16707. doi:10.1029/2009GL039640.

Hansen, J., M. Sato, P. Kharecha, D. Beerling, R. Berner, V. Masson-Delmotte, M. Pagani, M. Raymo, D. L. Royer, and J. C. Zachos (2008). "Target atmospheric CO2: Where should humanity aim?" The Open Atmospheric Science Journal 2: 217–231. doi:10.2174/1874282300802010217.

Hansen, J., M. Sato, P. Kharecha, G. Russell, D. W. Lea, and M. Siddall (2007). "Climate change and trace gases." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365(1856): 1925–1954. doi:10.1098/rsta.2007.2052.

Hansen, L. P. and T. J. Sargent (2001). "Acknowledging misspecification in macroeconomic theory." Review of Economic Dynamics 4(3): 519–535. doi:10.1006/redy.2001.0132.

Hart, R. (2008). "The timing of taxes on CO2 emissions when technological change is endogenous." Journal of Environmental Economics and Management 55(2): 194–212. doi:10.1016/j.jeem.2007.06.004.

Hastie, T., R. Tibshirani, and J. H. Friedman (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer series in statistics. New York: Springer, 2nd ed edition.

Hays, J. D., J. Imbrie, and N. J. Shackleton (1976). "Variations in the Earth's orbit: Pacemaker of the Ice Ages." Science 194(4270): 1121–1132. doi:10.1126/science.194.4270.1121.

Hegerl, G. C., T. J. Crowley, W. T. Hyde, and D. J. Frame (2006). "Climate sensitivity constrained by temperature reconstructions over the past seven centuries." Nature 440(7087): 1029–1032. doi:10.1038/nature04679.

Heimann, M. and M. Reichstein (2008). "Terrestrial ecosystem carbon dynamics and climate feedbacks." Nature 451(7176): 289–292. doi:10.1038/nature06591.

Helm, D., C. Hepburn, and R. Mash (2003). "Credible carbon policy." Oxford Review of Economic Policy 19(3): 438–450.

Helm, D., C. Hepburn, and R. Mash (2004). *Time Inconsistent Environmental Policy and Optimal Delegation*. Discussion Paper 175, Oxford University Department of Economics.

Hennlock, M. (2009). *Robust control in global warming management: An analytical dynamic integrated assessment*. Discussion paper 09-19, Resources For the Future.

Hepburn, C. (2006). "Regulation by prices, quantities, or both: A review of instrument choice." Oxford Review of Economic Policy 22(2): 226–247. doi:10.1093/oxrep/grj014.

Hitz, S. and J. Smith (2004). "Estimating global impacts from climate change." Global Environmental Change Part A 14(3): 201–218. doi:10.1016/j.gloenvcha.2004.04.010.

Hoffert, M. I., K. Caldeira, G. Benford, D. R. Criswell, C. Green, H. Herzog, A. K. Jain, H. S. Kheshgi, K. S. Lackner, J. S. Lewis, H. D. Lightfoot, W. Manheimer, J. C. Mankins, M. E. Mauel, L. J. Perkins, M. E. Schlesinger, T. Volk, and T. M. L. Wigley (2002). "Advanced technology paths to global climate stability: Energy for a greenhouse planet." Science 298(5595): 981–987. doi:10.1126/science.1072357.

Holdren, J. P. (2009). "The Administration's view of the state of the climate."

Hoogwijk, M., D. Vuuren, S. Boeters, K. Blok, E. Blomen, T. Barker, J. Chateau, A. Grübler, T. Masui, G. Nabuurs, A. Novikova, K. Riahi, S. R. du Can, J. Sathaye, S. Scrieciu, D. Urge-Vorsatz, and J. Vliet (2008). *Sectoral emission mitigation potentials: Comparing bottom-up and top-down approaches.* , Ecofys.

Huntingford, C., R. A. Fisher, L. Mercado, B. B. Booth, S. Sitch, P. P. Harris, P. M. Cox, C. D. Jones, R. A. Betts, Y. Malhi, G. R. Harris, M. Collins, and P. Moorcroft (2008). "Towards quantifying uncertainty in predictions of Amazon 'dieback'." Philosophical Transactions of the Royal Society B: Biological Sciences 363(1498): 1857 –1864. doi:10.1098/rstb.2007.0028.

Huntingford, C., J. A. Lowe, B. B. B. Booth, C. D. Jones, G. R. Harris, L. K. Gohar, and P. Meir (2009). "Contributions of carbon cycle uncertainty to future climate projection spread." Tellus B 61(2): 355–360. doi:10.1111/j.1600-0889.2009.00414.x.

Huybers, P. (2009). "Antarctica's orbital beat." Science 325(5944): 1085–1086. doi:10.1126/science.1176186.

Huybers, P. and G. Denton (2008). "Antarctic temperature at orbital timescales controlled by local summer duration." Nature Geoscience 1(11): 787–792. doi:10.1038/ngeo311.

Interagency Working Group on Social Cost of Carbon (2010). *Appendix 15a. Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866.* , United States Government.

Isik, M. (2004). "Incentives for technology adoption under environmental policy uncertainty: Implications for green payment programs." Environmental and Resource Economics 27(3): 247–263. doi:10.1023/B:EARE.0000017624.07757.3f.

Joos, F. and R. Spahni (2008). "Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years." Proceedings of the National Academy of Sciences 105(5): 1425 –1430. doi:10.1073/pnas.0707386105.

Jou, J. (2001). "Environment, asset characteristics, and optimal effluent fees." Environmental and Resource Economics 20(1): 27–39. doi:10.1023/A:1017524910959.

Jouzel, J., V. Masson-Delmotte, O. Cattani, G. Dreyfus, S. Falourd, G. Hoffmann, B. Minster, J. Nouet, J. M. Barnola, J. Chappellaz, H. Fischer, J. C. Gallet, S. Johnsen, M. Leuenberger, L. Loulergue, D. Luethi, H. Oerter, F. Parrenin, G. Raisbeck, D. Raynaud, A. Schilt, J. Schwander, E. Selmo, R. Souchez, R. Spahni, B. Stauffer, J. P. Steffensen, B. Stenni, T. F. Stocker, J. L. Tison, M. Werner, and E. W. Wolff (2007). "Orbital and millennial Antarctic climate variability over the past 800,000 years." Science 317(5839): 793–796. doi:10.1126/science.1141038.

Ju, N. and J. Miao (2009). *Ambiguity, learning, and asset returns.* Working Paper wp2009-014, Boston University.

Jun, M., R. Knutti, and D. W. Nychka (2008). "Spatial analysis to quantify numerical model bias and dependence." Journal of the American Statistical Association 103(483): 934–947. doi:10.1198/016214507000001265.

Kama, A. A. L., A. Pommeret, and F. Prieur (2011). *Optimal emission policy under the risk of irreversible pollution.* FEEM nota di lavoro 14.2011.

Karp, L. and Y. Tsur (2007). *Climate policy when the distant future matters: Catastrophic events with hyperbolic discounting.* CUDARE working paper 1037, University of California, Berkeley.

Karp, L. S. (2008). "Correct (and misleading) arguments for using market-based pollution control policies."

Kass, R. E. and L. Wasserman (1996). "The selection of prior distributions by formal rules." Journal of the American Statistical Association 91(435): 1343–1370. doi:10.2307/2291752.

Kawamura, K., F. Parrenin, L. Lisiecki, R. Uemura, F. Vimeux, J. P. Severinghaus, M. A. Hutterli, T. Nakazawa, S. Aoki, J. Jouzel, M. E. Raymo, K. Matsumoto, H. Nakata, H. Motoyama, S. Fujita, K. Goto-Azuma, Y. Fujii, and O. Watanabe (2007). "Northern Hemisphere forcing of climatic cycles in Antarctica over the past 360,000 years." Nature 448(7156): 912–916. doi:10.1038/nature06015.

Keith, D. W. (2009). "Why capture CO2 from the atmosphere?" Science 325(5948): 1654–1655. doi:10.1126/science.1175680.

Keith, D. W., M. Ha-Duong, and J. K. Stolaroff (2006). "Climate strategy with CO2 capture from the air." Climatic Change 74(1-3): 17–45. doi:10.1007/s10584-005-9026-x.

Keller, K., B. M. Bolker, and D. F. Bradford (2004). "Uncertain climate thresholds and optimal economic growth." Journal of Environmental Economics and Management 48(1): 723–741. doi:10.1016/j.jeem.2003.10.003.

Keller, K., G. Yohe, and M. Schlesinger (2008). "Managing the risks of climate thresholds: uncertainties and information needs." Climatic Change 91(1-2): 5–10. doi:10.1007/s10584-006-9114-6.

Kelly, D. L. and C. D. Kolstad (1999). "Bayesian learning, growth, and pollution." Journal of Economic Dynamics and Control 23(4): 491–518. doi:10.1016/S0165-1889(98)00034-7.

Kelly, D. L. and C. D. Kolstad (2001). "Solving infinite horizon growth models with an environmental sector." Computational Economics 18(2): 217–231. doi:10.1023/A:1021018417052.

Keynes, J. M. (1921). A Treatise on Probability. London: Macmillan and Co.

Kintisch, E. (2010). "'Asilomar 2' takes small steps toward rules for geoengineering." Science 328(5974): 22–23. doi:10.1126/science.328.5974.22.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). "A smooth model of decision making under ambiguity." Econometrica 73(6): 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). "Recursive smooth ambiguity preferences." Journal of Economic Theory 144(3): 930–976. doi:10.1016/j.jet.2008.10.007.

Knutti, R. (2008). "Should we believe model predictions of future climate change?" Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 366(1885): 4647–4664. doi:10.1098/rsta.2008.0169.

Knutti, R., M. R. Allen, P. Friedlingstein, J. M. Gregory, G. C. Hegerl, G. A. Meehl, M. Meinshausen, J. M. Murphy, G. Plattner, S. C. B. Raper, T. F. Stocker, P. A. Stott, H. Teng, and T. M. L. Wigley (2008). "A review of uncertainties in global temperature projections over the twenty-first century." Journal of Climate 21(11): 2651–2663. doi:10.1175/2007JCLI2119.1.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010). "Challenges in combining projections from multiple climate models." Journal of Climate 23(10): 2739–2758. doi:10.1175/2009JCLI3361.1.

Knutti, R. and G. C. Hegerl (2008). "The equilibrium sensitivity of the Earth's temperature to radiation changes." Nature Geoscience 1(11): 735–743. doi:10.1038/ngeo337.

Krey, V. and K. Riahi (2009). "Implications of delayed participation and technology failure for the feasibility, costs, and likelihood of staying below temperature targets—Greenhouse gas mitigation scenarios for the 21st century." Energy Economics 31(Supplement 2): S94–S106. doi:10.1016/j.eneco.2009.07.001.

Kriegler, E., J. W. Hall, H. Held, R. Dawson, and H. J. Schellnhuber (2009). “Imprecise probability assessment of tipping points in the climate system.” Proceedings of the National Academy of Sciences 106(13): 5041–5046. doi:10.1073/pnas.0809117106.

Lacis, A. A., G. A. Schmidt, D. Rind, and R. A. Ruedy (2010). “Atmospheric CO2: Principal control knob governing Earth’s temperature.” Science 330(6002): 356 –359. doi:10.1126/science.1190653.

Laffont, J. J. and J. Tirole (1996). “Pollution permits and environmental innovation.” Journal of Public Economics 62(1-2): 127–140. doi:10.1016/0047-2727(96)01576-9.

Lambert, P. C., A. J. Sutton, P. R. Burton, K. R. Abrams, and D. R. Jones (2005). “How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS.” Statistics in Medicine 24(15): 2401–2428. doi:10.1002/sim.2112.

Lange, A. and N. Treich (2008). “Uncertainty, learning and ambiguity in economic models on climate policy: some classical results and new directions.” Climatic Change 89(1): 7–21. doi:10.1007/s10584-008-9401-5.

Lawrence, D. M., A. G. Slater, R. A. Tomas, M. M. Holland, and C. Deser (2008). “Accelerated Arctic land warming and permafrost degradation during rapid sea ice loss.” Geophysical Research Letters 35: L11506. doi:10.1029/2008GL033985.

Le Quéré, C., C. Rodenbeck, E. T. Buitenhuis, T. J. Conway, R. Langenfelds, A. Gomez, C. Labuschagne, M. Ramonet, T. Nakazawa, N. Metzl, N. Gillett, and M. Heimann (2007). “Saturation of the Southern Ocean CO2 sink due to recent climate change.” Science 316(5832): 1735–1738. doi:10.1126/science.1136188.

Leach, A. J. (2007). “The climate change learning curve.” Journal of Economic Dynamics and Control 31(5): 1728–1752. doi:10.1016/j.jedc.2006.06.001.

Lehmann, J. (2007). “Bio-energy in the black.” Frontiers in Ecology and the Environment 5(7): 381–387. doi:10.1890/1540-9295(2007)5[381:BITB]2.0.CO;2.

Lemoine, D. M. (2010a). “Climate sensitivity distributions depend on the possibility that models share biases.” Journal of Climate 23(16): 4395–4415. doi:10.1175/2010JCLI3503.1.

Lemoine, D. M. (2010b). “Paleoclimatic warming increased carbon dioxide concentrations.” Journal of Geophysical Research 115: D22122. doi:10.1029/2010JD014725.

Lemoine, D. M., S. Fuss, J. Szolgayova, M. Obersteiner, and D. M. Kammen (2011). *The influence of negative emission technologies and technology policies on the optimal climate mitigation portfolio.* Working paper, University of California, Berkeley.

Lemoine, D. M. and C. Traeger (2010). *Tipping points and ambiguity in the integrated assessment of climate change*. CUDARE Working Paper 1111, University of California, Berkeley.

Lempert, R. J. and M. E. Schlesinger (2000). “Robust strategies for abating climate change.” Climatic Change 45(3): 387–401. doi:10.1023/A:1005698407365.

Lempert, R. J., M. E. Schlesinger, and J. K. Hammitt (1994). “The impact of potential abrupt climate changes on near-term policy choices.” Climatic Change 26(4): 351–376. doi:10.1007/BF01094402.

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.” Proceedings of the National Academy of Sciences 105(6): 1786–1793. doi:10.1073/pnas.0705414105.

Lenton, T. M. and N. E. Vaughan (2009). “The radiative forcing potential of different climate geoengineering options.” Atmospheric Chemistry and Physics Discussion 9(1): 2559–2608.

Luckow, P., M. Wise, J. Dooley, and S. Kim (2010). “Large-scale utilization of biomass energy and carbon dioxide capture and storage in the transport and electricity sectors under stringent CO2 concentration limit scenarios.” International Journal of Greenhouse Gas Control 4(5): 865–877. doi:10.1016/j.ijggc.2010.06.002.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter (2000). “WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility.” Statistics and Computing 10(4): 325–337. doi:10.1023/A:1008929526011.

Luo, Y. (2007). “Terrestrial carbon-cycle feedback to climate warming.” Annual Review of Ecology, Evolution, and Systematics 38(1): 683–712. doi:10.1146/annurev.ecolsys.38.091206.095808.

Lüthi, D., M. L. Floch, B. Bereiter, T. Blunier, J. Barnola, U. Siegenthaler, D. Raynaud, J. Jouzel, H. Fischer, K. Kawamura, and T. F. Stocker (2008). “High-resolution carbon dioxide concentration record 650,000-800,000 years before present.” Nature 453(7193): 379–382. doi:10.1038/nature06949.

MacFarling Meure, C., D. Etheridge, C. Trudinger, P. Steele, R. Langenfelds, T. van Ommen, A. Smith, and J. Elkins (2006). “Law Dome CO2, CH4 and N2O ice core records extended to 2000 years BP.” Geophysical Research Letters 33: L14810. doi:10.1029/2006GL026152.

Malik, A. S. (1991). “Permanent versus interim regulations: A game-theoretic analysis.” Journal of Environmental Economics and Management 21(2): 127–139. doi:10.1016/0095-0696(91)90037-J.

Mann, M. E., Z. Zhang, M. K. Hughes, R. S. Bradley, S. K. Miller, S. Rutherford, and F. Ni (2008). "Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia." Proceedings of the National Academy of Sciences 105(36): 13252–13257. doi:10.1073/pnas.0805721105.

Masson-Delmotte, V., B. Stenni, K. Pol, P. Braconnot, O. Cattani, S. Falourd, M. Kageyama, J. Jouzel, A. Landais, B. Minster, J. Barnola, J. Chappellaz, G. Krinner, S. Johnsen, R. Röthlisberger, J. Hansen, U. Mikolajewicz, and B. Otto-Bliesner (2010). "EPICA Dome C record of glacial and interglacial intensities." Quaternary Science Reviews 29(1-2): 113–128. doi:10.1016/j.quascirev.2009.09.030.

Masur, J. S. and E. A. Posner (2010). *Climate regulation and the limits of cost-benefit analysis*. John m. olin law & economics working paper no. 525, The Law School, University of Chicago.

Matthews, H. D., M. Eby, T. Ewen, P. Friedlingstein, and B. J. Hawkins (2007). "What determines the magnitude of carbon cycle-climate feedbacks?" Global Biogeochemical Cycles 21: GB2012. doi:10.1029/2006GB002733.

Matthews, H. D., N. P. Gillett, P. A. Stott, and K. Zickfeld (2009). "The proportionality of global warming to cumulative carbon emissions." Nature 459(7248): 829–832. doi: 10.1038/nature08047.

McJeon, H. C., L. Clarke, P. Kyle, M. Wise, A. Hackbarth, B. P. Bryant, and R. J. Lempert (in press). "Technology interactions among low-carbon energy technologies: What can we learn from a large number of scenarios?" Energy Economics doi:10.1016/j.eneco.2010.10.007.

Meehl, G., T. F. Stocker, W. D. Collins, P. Friedlingstein, A. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, S. C. B. Raper, I. Watterson, A. J. Weaver, and Z. Zhao (2007). "Global climate projections." In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, eds., "Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," pp. 748–845. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

MEF (2009). "Declaration of the Leaders: The Major Economies Forum on Energy and Climate." L'Aquila, Italy.

Meinshausen, M., N. Meinshausen, W. Hare, S. C. B. Raper, K. Frieler, R. Knutti, D. J. Frame, and M. R. Allen (2009). "Greenhouse-gas emission targets for limiting global warming to 2°C." Nature 458(7242): 1158–1162. doi:10.1038/nature08017.

Milanković, M. (1941). Kanon der Erdbestrahlung und seine Andwendung auf das Eiszeitenproblem. Belgrade, Serbia: Royal Serbian Academy.

Millner, A., S. Dietz, and G. Heal (2010). “Ambiguity and climate policy.” National Bureau of Economic Research Working Paper Series No. 16050.

Min, S. and A. Hense (2007). “Hierarchical evaluation of IPCC AR4 coupled climate models with systematic consideration of model uncertainties.” Climate Dynamics 29(7): 853–868. doi:10.1007/s00382-007-0269-2.

Miranda, M. J. and P. L. Fackler (2002). Applied Computational Economics and Finance. Cambridge, Massachusetts: MIT Press.

Montgomery, W. D. and A. E. Smith (2007). “Price, quantity, and technology strategies for climate change policy.” In M. E. Schlesinger, H. S. Kheshgi, J. Smith, F. C. de la Chesnaye, J. M. Reilly, T. Wilson, and C. Kolstad, eds., “Human-Induced Climate Change: An interdisciplinary assessment,” pp. 328–342. New York: Cambridge University Press.

Nævdal, E. (2006). “Dynamic optimisation in the presence of threshold effects when the location of the threshold is uncertain—with an application to a possible disintegration of the Western Antarctic Ice Sheet.” Journal of Economic Dynamics and Control 30(7): 1131–1158. doi:10.1016/j.jedc.2005.04.004.

Nævdal, E. and M. Oppenheimer (2007). “The economics of the thermohaline circulation—A problem with multiple thresholds of unknown locations.” Resource and Energy Economics 29(4): 262–283. doi:10.1016/j.reseneeco.2007.01.003.

National Research Council (2011). Climate Stabilization Targets: Emissions, Concentrations, and Impacts over Decades to Millennia. Washington, D.C.: The National Academies Press.

Nemet, G. F. (2006). “Beyond the learning curve: factors influencing cost reductions in photovoltaics.” Energy Policy 34(17): 3218–3232. doi:10.1016/j.enpol.2005.06.020.

Newbold, S. C. and A. Daigneault (2009). “Climate response uncertainty and the benefits of greenhouse gas emissions reductions.” Environmental and Resource Economics 44(3): 351–377. doi:10.1007/s10640-009-9290-8.

Nordhaus, W. D. (1992). “An optimal transition path for controlling greenhouse gases.” Science 258(5086): 1315–1319.

Nordhaus, W. D. (1993). “Rolling the ‘DICE’: an optimal transition path for controlling greenhouse gases.” Resource and Energy Economics 15(1): 27–50. doi:10.1016/0928-7655(93)90017-O.

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

Nordhaus, W. D. and J. Boyer (2000). Warming the World: Economic Models of Global Warming. Cambridge, Mass: MIT Press.

Notz, D. (2009). "The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss." Proceedings of the National Academy of Sciences 106(49): 20590 – 20595. doi:10.1073/pnas.0902356106.

Obersteiner, M., C. Azar, P. Kauppi, K. Mollersten, J. Moreira, S. Nilsson, P. Read, K. Riahi, B. Schlamadinger, Y. Yamagata, J. Yan, and J. van Ypersele (2001). "Managing climate risk." Science 294(5543): 786b–787. doi:10.1126/science.294.5543.786b.

O'Neill, B. C. and M. Oppenheimer (2004). "Climate change impacts are sensitive to the concentration stabilization path." Proceedings of the National Academy of Sciences of the United States of America 101(47): 16411–16416. doi:10.1073/pnas.0405522101.

Oppenheimer, M. (1998). "Global warming and the stability of the West Antarctic Ice Sheet." Nature 393(6683): 325–332. doi:10.1038/30661.

Oreskes, N. and E. M. Conway (2010). Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming. New York: Bloomsbury Press.

Overpeck, J. T. and J. E. Cole (2006). "Abrupt change in Earth's climate system." Annual Review of Environment and Resources 31(1): 1–31. doi:10.1146/annurev.energy.30.050504.144308.

Parrenin, F., F. Rémy, C. Ritz, M. J. Siegert, and J. Jouzel (2004). "New modeling of the Vostok ice flow line and implication for the glaciological chronology of the Vostok ice core." Journal of Geophysical Research Atmospheres 109: D20102. doi:10.1029/2004JD004561.

Petit, J. R., J. Jouzel, D. Raynaud, N. I. Barkov, J. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davis, G. Delaygue, M. Delmotte, V. M. Kotlyakov, M. Legrand, V. Y. Lipenkov, C. Lorius, L. Pépin, C. Ritz, E. Saltzman, and M. Stievenard (1999). "Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica." Nature 399(6735): 429–436. doi:10.1038/20859.

Pielke Jr., R. A. (2009). "An idealized assessment of the economics of air capture of carbon dioxide in mitigation policy." Environmental Science & Policy 12(3): 216–225. doi:10.1016/j.envsci.2009.01.002.

Pierrehumbert, R. T. (2011). "Infrared radiation and planetary temperature." Physics Today 64(1): 33. doi:10.1063/1.3541943.

Pindyck, R. S. (2000). “Irreversibilities and the timing of environmental policy.” Resource and Energy Economics 22(3): 233–259. doi:10.1016/S0928-7655(00)00033-6.

Pindyck, R. S. (2002). “Optimal timing problems in environmental economics.” Journal of Economic Dynamics & Control 26(9-10): 1677–1697. doi:10.1016/S0165-1889(01)00090-2.

Pindyck, R. S. (2007). “Uncertainty in environmental economics.” Review of Environmental Economics and Policy 1(1): 45–65. doi:10.1093/reep/rem002.

Pizer, W. A. and D. Popp (2008). “Endogenizing technological change: Matching empirical evidence to modeling needs.” Energy Economics 30(6): 2754–2770. doi:10.1016/j.eneco.2008.02.006.

Plattner, G., R. Knutti, F. Joos, T. F. Stocker, W. von Bloh, V. Brovkin, D. Cameron, E. Driesschaert, S. Dutkiewicz, M. Eby, N. R. Edwards, T. Fichefet, J. C. Hargreaves, C. D. Jones, M. F. Loutre, H. D. Matthews, A. Mouchet, S. A. Müller, S. Nawrath, A. Price, A. Sokolov, K. M. Strassmann, and A. J. Weaver (2008). “Long-term climate commitments projected with climate-carbon cycle models.” Journal of Climate 21(12): 2721–2751. doi:10.1175/2007JCLI1905.1.

Plevin, R. J. (2010). *Life cycle regulation of transportation fuels: Uncertainty and its policy implications*. Dissertation, Energy and Resources Group, University of California, Berkeley.

Polasky, S., A. de Zeeuw, and F. Wagener (2010). *Optimal management with potential regime shifts*. Tinbergen Institute Discussion Paper TI 2010-111/1.

Popp, D. (2010). “Innovation and climate policy.” Annual Review of Resource Economics 2(1): 275–298. doi:10.1146/annurev.resource.012809.103929.

Popp, D., R. G. Newell, and A. B. Jaffe (2009). “Energy, the environment, and technological change.” National Bureau of Economic Research Working Paper Series No. 14832.

Ramanathan, V. and Y. Feng (2008). “On avoiding dangerous anthropogenic interference with the climate system: Formidable challenges ahead.” Proceedings of the National Academy of Sciences 105(38): 14245 –14250. doi:10.1073/pnas.0803838105.

Ramaswamy, V., O. Boucher, J. Haigh, D. Hauglustaine, J. Haywood, G. Myhre, T. Nakajima, G. Shi, and S. Solomon (2001). “Radiative forcing of climate change.” In J. Houghton, Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. Johnson, eds., “Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change,” p. 881. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Raupach, M. R., J. G. Canadell, and C. Le Quéré (2008). "Anthropogenic and biophysical contributions to increasing atmospheric CO2 growth rate and airborne fraction." Biogeosciences 5: 1601–1613. doi:10.5194/bg-5-1601-2008.

Read, P. (2009). "Reducing CO2 levels—so many ways, so few being taken." Climatic Change 97(3-4): 449–458. doi:10.1007/s10584-009-9723-y.

Requate, T. (2005). "Dynamic incentives by environmental policy instruments—a survey." Ecological Economics 54(2-3): 175–195. doi:10.1016/j.ecolecon.2004.12.028.

Requate, T. and W. Unold (2003). "Environmental policy incentives to adopt advanced abatement technology: Will the true ranking please stand up?" European Economic Review 47(1): 125–146. doi:10.1016/S0014-2921(02)00188-5.

Rhodes, J. S. and D. W. Keith (2005). "Engineering economic analysis of biomass IGCC with carbon capture and storage." Biomass and Bioenergy 29(6): 440–450. doi:10.1016/j.biombioe.2005.06.007.

Riahi, K., A. Grübler, and N. Nakićenović (2007). "Scenarios of long-term socio-economic and environmental development under climate stabilization." Technological Forecasting and Social Change 74(7): 887–935. doi:10.1016/j.techfore.2006.05.026.

Rockafellar, R. T. and S. Uryasev (2000). "Optimization of conditional value-at-risk." Journal of Risk 2(3): 21–41.

Rockström, J., W. Steffen, K. Noone, Åsa Persson, F. S. Chapin, E. F. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. A. Foley (2009). "A safe operating space for humanity." Nature 461(7263): 472–475. doi:10.1038/461472a.

Roe, G. H. (2009). "Feedbacks, timescales, and seeing red." Annual Review of Earth and Planetary Sciences 37(1): 93–115. doi:10.1146/annurev.earth.061008.134734.

Roe, G. H. and M. B. Baker (2007). "Why is climate sensitivity so unpredictable?" Science 318(5850): 629–632. doi:10.1126/science.1144735.

Salamatin, A. N. (2000). "Paleoclimatic reconstructions based on borehole temperature measurements in ice sheets. Possibilities and limitations." In "Physics of Ice Core Records," pp. 243–282. Shikotsukohan, Hokkaido, Japan: Hokkaido University Press.

Sandén, B. A. and C. Azar (2005). "Near-term technology policies for long-term climate targets—economy wide versus technology specific approaches." Energy Policy 33(12): 1557–1576. doi:10.1016/j.enpol.2004.01.012.

Schaefer, K., T. Zhang, L. Bruhwiler, and A. P. Barrett (2011). “Amount and timing of permafrost carbon release in response to climate warming.” Tellus B 63(2): 165–180. doi:10.1111/j.1600-0889.2011.00527.x.

Scheffer, M., V. Brovkin, and P. M. Cox (2006). “Positive feedback between global warming and atmospheric CO2 concentration inferred from past climate change.” Geophysical Research Letters 33: L10702. doi:10.1029/2005GL025044.

Schimel, D. S., B. H. Braswell, R. McKeown, D. S. Ojima, W. J. Parton, and W. Pulliam (1996). “Climate and nitrogen controls on the geography and timescales of terrestrial biogeochemical cycling.” Global Biogeochemical Cycles 10(4): 677–692. doi:10.1029/96GB01524.

Schmidt, G. A. and D. T. Shindell (2003). “Atmospheric composition, radiative forcing, and climate change as a consequence of a massive methane release from gas hydrates.” Paleoceanography 18(1): 1004. doi:200310.1029/2002PA000757.

Schuur, E. A. G., J. Bockheim, J. G. Canadell, E. Euskirchen, C. B. Field, S. V. Goryachkin, S. Hagemann, P. Kuhry, P. M. Lafleur, H. Lee, G. Mazhitova, F. E. Nelson, A. Rinke, V. E. Romanovsky, N. Shiklomanov, C. Tarnocai, S. Venevsky, J. G. Vogel, and S. A. Zimov (2008). “Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle.” BioScience 58(8): 701–714. doi:10.1641/B580807.

Schuur, E. A. G., J. G. Vogel, K. G. Crummer, H. Lee, J. O. Sickman, and T. E. Osterkamp (2009). “The effect of permafrost thaw on old carbon release and net carbon exchange from tundra.” Nature 459(7246): 556–559. doi:10.1038/nature08031.

Sitch, S., C. Huntingford, N. Gedney, P. E. Levy, M. Lomas, S. L. Piao, R. Betts, P. Ciais, P. Cox, P. Friedlingstein, C. D. Jones, I. C. Prentice, and F. I. Woodward (2008). “Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs).” Global Change Biology 14(9): 2015–2039. doi:10.1111/j.1365-2486.2008.01626.x.

Smetacek, V. and S. Naqvi (2008). “The next generation of iron fertilization experiments in the Southern Ocean.” Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 366(1882): 3947–3967. doi:10.1098/rsta.2008.0144.

Smith, J. B., S. H. Schneider, M. Oppenheimer, G. W. Yohe, W. Hare, M. D. Mastrandrea, A. Patwardhan, I. Burton, J. Corfee-Morlot, C. H. D. Magadza, H. Fussel, A. B. Pittock, A. Rahman, A. Suarez, and J. van Ypersele (2009). “Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) “reasons for concern”.” Proceedings of the National Academy of Sciences 106(11): 4133–4137. doi:10.1073/pnas.0812355106.

Soden, B. J. and I. M. Held (2006). "An assessment of climate feedbacks in coupled ocean-atmosphere models." Journal of Climate 19(14): 3354–3360. doi:10.1175/JCLI3799.1.

Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields (2008). "Quantifying climate feedbacks using radiative kernels." Journal of Climate 21(14): 3504–3520. doi:10.1175/2007JCLI2110.1.

Sokolov, A. P., D. W. Kicklighter, J. M. Melillo, B. S. Felzer, C. A. Schlosser, and T. W. Cronin (2008). "Consequences of considering carbon-nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle." Journal of Climate 21(15): 3776–3796. doi:10.1175/2008JCLI2038.1.

Solomon, S., G. Plattner, R. Knutti, and P. Friedlingstein (2009). "Irreversible climate change due to carbon dioxide emissions." Proceedings of the National Academy of Sciences 106(6): 1704–1709. doi:10.1073/pnas.0812721106.

Spahni, R., J. Schwander, J. Flückiger, B. Stauffer, J. Chappellaz, and D. Raynaud (2003). "The attenuation of fast atmospheric CH4 variations recorded in polar ice cores." Geophysical Research Letters 30(11): 1571. doi:200310.1029/2003GL017093.

Spencer, R. W. and W. D. Braswell (2008). "Potential biases in feedback diagnosis from observational data: A simple model demonstration." Journal of Climate 21(21): 5624–5628. doi:10.1175/2008JCLI2253.1.

Steinhilber, F., J. Beer, and C. Fröhlich (2009). "Total solar irradiance during the Holocene." Geophysical Research Letters 36: L19704. doi:10.1029/2009GL040142.

Stephens, G. L. (2005). "Cloud feedbacks in the climate system: A critical review." Journal of Climate 18(2): 237–273. doi:10.1175/JCLI-3243.1.

Stephens, J. and D. Keith (2008). "Assessing geochemical carbon management." Climatic Change 90(3): 217–242. doi:10.1007/s10584-008-9440-y.

Stolaroff, J. K., D. W. Keith, and G. V. Lowry (2008). "Carbon dioxide capture from atmospheric air using sodium hydroxide spray." Environmental Science & Technology 42(8): 2728–2735. doi:10.1021/es702607w.

Strand, S. E. and G. Benford (2009). "Ocean sequestration of crop residue carbon: Recycling fossil fuel carbon back to deep sediments." Environmental Science & Technology 43(4): 1000–1007. doi:10.1021/es8015556.

Strong, A. L., J. J. Cullen, and S. W. Chisholm (2009). "Ocean fertilization: Science, policy, and commerce." Oceanography 22(3): 236–261.

Tarnocai, C., J. G. Canadell, E. A. G. Schuur, P. Kuhry, G. Mazhitova, and S. Zimov (2009). "Soil organic carbon pools in the northern circumpolar permafrost region." Global Biogeochemical Cycles 23: GB2023. doi:10.1029/2008GB003327.

Tarui, N. and S. Polasky (2005). "Environmental regulation with technology adoption, learning and strategic behavior." Journal of Environmental Economics and Management 50(3): 447–467. doi:10.1016/j.jeem.2005.01.004.

Tebaldi, C. and R. Knutti (2007). "The use of the multi-model ensemble in probabilistic climate projections." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365(1857): 2053–2075. doi:10.1098/rsta.2007.2076.

Tebaldi, C. and B. Sansó (2009). "Joint projections of temperature and precipitation change from multiple climate models: a hierarchical Bayesian approach." Journal of the Royal Statistical Society: Series A (Statistics in Society) 172(1): 83–106. doi: {10.1111/j.1467-985X.2008.00545.x}.

Tol, R. S. (2008). "The social cost of carbon: Trends, outliers and catastrophes." Economics: The Open-Access, Open Assessment E-Journal 2(2008-25).

Tol, R. S. J. (2009). "The economic effects of climate change." Journal of Economic Perspectives 23(2): 29–51. doi:10.1257/jep.23.2.29.

Tomassini, L., P. Reichert, R. Knutti, T. F. Stocker, and M. E. Borsuk (2007). "Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods." Journal of Climate 20(7): 1239–1254. doi:10.1175/JCLI4064.1.

Torn, M. S. and J. Harte (2006). "Missing feedbacks, asymmetric uncertainties, and the underestimation of future warming." Geophysical Research Letters 33: L10703. doi: 10.1029/2005GL025540.

Traeger, C. P. (2010a). *Intertemporal risk aversion—or—Wouldn't it be nice to tell whether Robinson Crusoe is risk averse?* CUDARE Working Paper 1102, University of California, Berkeley.

Traeger, C. P. (2010b). *Subjective risk, confidence, and ambiguity.* CUDARE Working Paper 1103, University of California, Berkeley.

Tsur, Y. and A. Zemel (1996). "Accounting for global warming risks: Resource management under event uncertainty." Journal of Economic Dynamics and Control 20(6-7): 1289–1305. doi:10.1016/0165-1889(95)00900-0.

Uddin, S. N. and L. Barreto (2007). "Biomass-fired cogeneration systems with CO2 capture and storage." Renewable Energy 32(6): 1006–1019. doi:10.1016/j.renene.2006.04.009.

UNFCCC (2009). "Copenhagen Accord."

UNFCCC (2011). "Compilation of economy-wide emission reduction targets to be implemented by Parties included in Annex I to the Convention: Note by the secretariat."

Urban, N. M. and K. Keller (2009). "Complementary observational constraints on climate sensitivity." Geophysical Research Letters 36: L04708. doi:10.1029/2008GL036457.

van der Zwaan, B. and R. Gerlagh (2009). "Economics of geological CO2 storage and leakage." Climatic Change 93(3): 285–309. doi:10.1007/s10584-009-9558-6.

van Soest, D. P. (2005). "The impact of environmental policy instruments on the timing of adoption of energy-saving technologies." Resource and Energy Economics 27(3): 235–247. doi:10.1016/j.reseneeco.2004.11.002.

van Vuuren, D., M. den Elzen, P. Lucas, B. Eickhout, B. Strengers, B. van Ruijven, S. Wonink, and R. van Houdt (2007). "Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs." Climatic Change 81(2): 119–159. doi:10.1007/s10584-006-9172-9.

van Vuuren, D. P., E. Bellevrat, A. Kitous, and M. Isaac (2010a). "Bio-energy use and low stabilization scenarios." Energy Journal 31: 193–221.

van Vuuren, D. P. and K. Riahi (2011). "The relationship between short-term emissions and long-term concentration targets." Climatic Change 104(3-4): 793–801. doi:10.1007/s10584-010-0004-6.

van Vuuren, D. P., E. Stehfest, M. G. den Elzen, J. van Vliet, and M. Isaac (2010b). "Exploring IMAGE model scenarios that keep greenhouse gas radiative forcing below 3 W/m2 in 2100." Energy Economics 32(5): 1105–1120. doi:10.1016/j.eneco.2010.03.001.

van Vuuren, D. P., J. van Vliet, and E. Stehfest (2009). "Future bio-energy potential under various natural constraints." Energy Policy 37(11): 4220–4230. doi:10.1016/j.enpol.2009.05.029.

Vaughan, D. G. (2008). "West Antarctic Ice Sheet collapse—the fall and rise of a paradigm." Climatic Change 91(1-2): 65–79. doi:10.1007/s10584-008-9448-3.

Vissing-Jørgensen, A. and O. P. Attanasio (2003). "Stock-market participation, intertemporal substitution, and risk-aversion." The American Economic Review 93(2): 383–391.

von Döllen, A. and T. Requate (2008). "Environmental policy and uncertain arrival of future abatement technology." The B.E. Journal of Economic Analysis & Policy 8(1): Article 30. doi:10.2202/1935-1682.1951.

von Neumann, J. and O. Morgenstern (1944). Theory of Games and Economic Behavior. Princeton: Princeton University Press.

Warren, R., M. Mastrandrea, C. Hope, and A. Hof (2010). "Variation in the climatic response to SRES emissions scenarios in integrated assessment models." Climatic Change 102(3): 671–685. doi:10.1007/s10584-009-9769-x.

Weber, T. A. and K. Neuhoff (2010). "Carbon markets and technological innovation." Journal of Environmental Economics and Management 60(2): 115–132. doi:10.1016/j.jeem.2010.04.004.

Weil, P. (1990). "Nonexpected utility in macroeconomics." The Quarterly Journal of Economics 105(1): 29–42.

Weitzman, M. L. (1974). "Prices vs. quantities." The Review of Economic Studies 41(4): 477–491. doi:10.2307/2296698.

Weitzman, M. L. (2009). "On modeling and interpreting the economics of catastrophic climate change." Review of Economics and Statistics 91(1): 1–19. doi:10.1162/rest.91.1.1.

Williams, P. D., E. Guilyardi, R. Sutton, J. Gregory, and G. Madec (2007). "A new feedback on climate change from the hydrological cycle." Geophysical Research Letters 34: L08706. doi:10.1029/2007GL029275.

Winton, M. (2006). "Surface albedo feedback estimates for the AR4 climate models." Journal of Climate 19(3): 359–365. doi:10.1175/JCLI3624.1.

Wolff, E. W., H. Fischer, and R. Röthlisberger (2009). "Glacial terminations as southern warmings without northern control." Nature Geoscience 2(3): 206–209. doi:10.1038/ngeo442.

Woodward, F. I., R. D. Bardgett, J. A. Raven, and A. M. Hetherington (2009). "Biological approaches to global environment change mitigation and remediation." Current Biology 19(14): R615–R623. doi:10.1016/j.cub.2009.06.012.

Wright, E. L. and J. D. Erickson (2003). "Incorporating catastrophes into integrated assessment: Science, impacts, and adaptation." Climatic Change 57(3): 265–286. doi:10.1023/A:1022829706609.

Yoshimori, M., T. Yokohata, and A. Abe-Ouchi (2009). "A comparison of climate feedback strength between CO2 doubling and LGM experiments." Journal of Climate 22(12): 3374–3395. doi:10.1175/2009JCLI2801.1.

Zachos, J. C., G. R. Dickens, and R. E. Zeebe (2008). "An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics." Nature 451(7176): 279–283. doi: 10.1038/nature06588.

Zaliapin, I. and M. Ghil (2010). "Another look at climate sensitivity." Nonlinear Processes in Geophysics 17(2): 113–122. doi:10.5194/npg-17-113-2010.

Zeebe, R. E., J. C. Zachos, and G. R. Dickens (2009). "Carbon dioxide forcing alone insufficient to explain Palaeocene-Eocene Thermal Maximum warming." Nature Geoscience 2: 576–580. doi:10.1038/ngeo578.

Zhao, J. (2003). "Irreversible abatement investment under cost uncertainties: tradable emission permits and emissions charges." Journal of Public Economics 87(12): 2765–2789. doi:10.1016/S0047-2727(02)00135-4.