USING INTEGRATED ASSESSMENT MODELS TO ACHIEVE THE PARIS CLIMATE TARGET

Ida Andrea Braathen Sognnæs

Department of Land Economy

University of Cambridge

This thesis is submitted for the degree of

Doctor of Philosophy

Emmanuel College December 2019

Declaration

This thesis is submitted according to the requirements of the Degree Committee of Land Economy. It does not exceed the regulation length of 80,000 words including footnotes, references, and appendices. It is the result of my own work and includes nothing which is the outcome of work done in collaboration with others, except where specifically indicated in the text and Acknowledgements.

Ida Andrea Braathen Sognnæs

Cambridge, December 2019

Abstract

Thesis title: Using Integrated Assessment Models to Achieve the Paris Climate Target

Author: Ida Andrea Braathen Sognnæs

Integrated assessment models (IAMs) have become central tools in global assessments of how to achieve the Paris climate target. But how reliable are the insights that can be drawn from IAMs? This thesis identifies and begins to assess three challenges associated with the use of IAM ensembles and individual IAMs to draw insights on climate mitigation.

First, it highlights the importance of model independence for the robustness of insights that can be drawn from IAM ensembles. It develops a method that uses model documentation to construct a model family tree and uses this method to identify likely model dependencies between IAMs in the IPCC's 5th assessment report (AR5). The analysis shows that the 14 most influential IAMs in AR5 form three branches, the largest of which (including MERGE, MESSAGE, MERGE-ETL, REMIND, WITCH, and BET) is responsible for about half of the scenarios in AR5. The model documentation also indicates that an expanding set of policy questions has incentivised a continuous increase in the detail and scope of IAMs over time. The findings give reason to believe that the diversity of model choices and assumptions included in the AR5 IAM ensemble might be limited.

Second, it argues, based on a debate on values in science in philosophy, that the exclusively positive estimates of the cost of mitigation in AR5 are problematic because they don't capture the full range of cost estimates found in the literature and because the uncertainty of the cost of mitigation is important. A review of the literature reveals that general equilibrium models, which are responsible for all the cost estimates in AR5, can (despite claims to the contrary) be modified to generate net negative costs, but that few of the IAMs in AR5 include mechanisms that typically contribute to net negative costs. It is also found that the model intercomparison studies that are responsible for most of the AR5 cost estimates focused only on aspects that increase the cost of mitigation. Overall, this gives reason to believe that the AR5 IAM ensemble might be biased against net negative mitigation costs.

Third, it shows that predictions of climate policy impacts based on the Future Technology Transformations (FTT) simulation model are highly sensitive to a scaling parameter whose correct value is deeply uncertain. This result, which is obtained using Monte Carlo analysis and uniform and independent distributions (around ±50% of default values) for investor discount rates, technology build times, technology lifetimes, learning rates, and the scaling factor in a global sensitivity analysis, shows that the use of diffusion theory to derive technology deployment – which is seen by those who designed FTT to present a unique feature of the model – does not in itself ensure reliable predictions. In fact, the result indicates that predictions from both energy system optimisation models, which are more widely used, and FTT depend on similar unknowns related to future rates of technological change.

Based on these three challenges, the thesis concludes, a diversity of model choices and assumptions is crucial for ensuring robust insights and for reflecting important uncertainties associated with IAM research.

Acknowledgments

First and foremost, I would like to thank my husband, Jack Wright. This thesis would not have been possible without his love and support. Our countless conversations and hours spent pondering questions, both small and large, have had a profound impact on the way I think about the topics that are covered in this thesis and beyond. Jack has also provided helpful and detailed comments on the final draft, while at the same time keeping the house in order and taking care of our son during the final stretches. I would also like to thank our son, Leon Sognnæs Wright, for bringing a constant stream of smiles and laughter into my life during my doctoral research.

I would like to thank my supervisor, Jonathan Aldred, for always making me think, and for his detailed comments on the final draft. I would also like to thank Jean-Francois Mercure for numerous interesting conversations and Pablo Salas for sharing his many insights on modelling.

Beyond these individuals, there are several groups of people in Cambridge that have inspired and challenged my thinking during my PhD. I would like to thank everyone involved in setting up and running "Paper 0: Neglects of Neoclassical Economics", the "Politics of Economics" research network, members of Ha-Joon Chang's reading group, and the Cambridge Social Ontology Group. I would also like to thank the Cambridge Land Economy Advisory Board for supporting my doctoral research through three years, the In Search of 'Good' Energy Policy Research group, for providing me with support and a small grant, and the Cambridge Political Economy Society Trust for funding the last six months of my thesis writing. I would also like to thank Rupert Way and Doyne Farmer for hosting me while visiting the Institute for New Economic Thinking (INET) in Oxford.

Finally, I would like to thank my friends and family around the world, especially Jessica Joan Goddard, with whom I've shared a long intellectual journey in between traditional disciplines. Lastly, I would like to thank my mum, Johanne Sognnæs, for her support and excitement at the outset of this journey. Has she still been around, I know she would have been proud to see the end result.

Table of Contents

1 Introduction

The Paris Agreement on climate change was adopted by consensus at the 21st Conference of the Parties to the United Nation Framework Convention on Climate Change on 12 December 2015. The central aim of the agreement is to "strengthen the global response to the threat of climate change by keeping a global temperature rise this century well below 2 degrees Celsius above pre-industrial levels and to pursue efforts to limit the temperature increase even further to 1.5 degrees Celsius" (UNFCCC, 2019).

There is now widespread agreement that global warming is caused by anthropogenic emissions of greenhouse gases (GHGs) (IPCC, 2014b). According to the Intergovernmental Panel on Climate Change (IPCC) *Special Report on the impacts of global warming of 1.5°C* (SR15) (IPCC, 2018a), GHG emissions have already raised global temperatures by roughly 1.0˚C since pre-industrial times. Global CO_2 emissions now need to reach net zero around 2050 if we want to limit global warming to below 1.5˚C and around 2070 if we want to limit global warming to below 2˚C[1]. Given that CO_2 emissions have continued to rise, with 2018 seeing record levels (Jackson et al., 2018), meeting the Paris climate target will require a radical shift from current trends. Figure 1.1 provides a graphical depiction of what is required. According to Rockström et al. (2017), this will take "herculean efforts" and according to Stern (2009) "we will need a revolution that surpasses the scale and impact of previous world-changing technologies" in order to get there. What might this revolution look like?

[1] Reaching net zero by 2050 gives a 50% chance of staying below 1.5˚C and reaching net zero by 2070 gives a 66% chance of staying below 2˚C. In these scenarios the emissions of non-CO_2 GHGs are also reduced, but they do not reach zero.

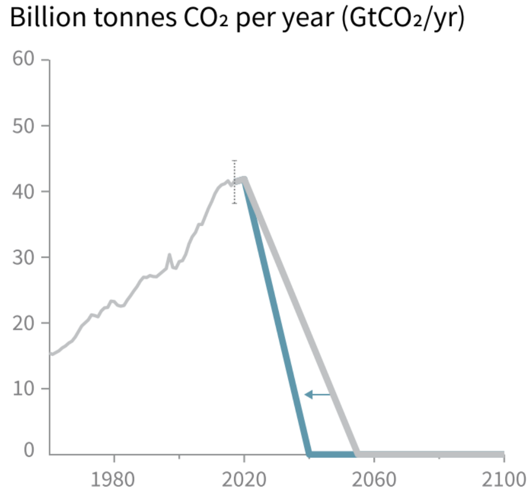


Figure 1.1 Stylized net global CO_2 emission pathways for 1.5˚C (the blue pathway gives to a higher probability that temperatures will stay below 1.5˚C than the grey pathway). Reproduced from IPCC (2018b, p. 8).

Integrated Assessment Models (IAMs) have come to play a crucial role in providing answers to this question. By computing different scenarios, IAMs tell us how human systems (most notably the energy system and the economy) might evolve so as to achieve different temperature targets. More specifically, a *scenario* is defined in the IPCC's 5th assessment report (AR5) as

> "a plausible description of how the future may develop based on a coherent and internally consistent set of assumptions about key driving forces (e.g., rate of technological change (TC), prices) and relationships." (IPCC, 2014a, p. 1270).

Similarly, a scenario can also be thought of as "a story of what happened in the future" (Knutti, 2018, p. 214). Essentially, in AR5, *transformation pathways* are *scenarios* in which emissions are reduced[2]. In concrete terms, a scenario in AR5 is equivalent to a set of (quantitative) IAM outputs such as the share of low-carbon energy, electricity demand, emissions by sector, natural resource use, land use, and GDP. Together, such IAM outputs can be used to describe how the world might evolve to achieve the Paris climate target.

Some of the most well-known IAMs, such as DICE (Nordhaus, 1992) and PAGE (Hope et al., 2003), incorporate the damages from climate change and endogenously balance damages with mitigation costs. These IAMs, which might be referred to as cost benefit analysis IAMs (or CBA IAMs), tend to be more stylized and are not the main focus in the IPCC's most recent reports. Instead, both AR5 and SR15 rely on larger scale non-CBA IAMs (referred to here as large-scale IAMs or simply IAMs) to generate long-term scenarios for climate mitigation and to assess different transformation pathways. These IAMs are preferred partly because CBA IAMs tend to be highly aggregated and according to AR5 don't provide

[2] There are also baseline scenarios, in which nothing is (intentionally) 'transformed'.

the "sufficient sectoral and geographic resolution to understand the evolution of key processes such as energy systems or land systems" (2014a, p. 51). In addition, CBA IAMs have received a significant amount of criticism, in particular for the way in which they use damage functions and social discounting (e.g. Pindyck, 2013; Stern, 2013b; Weitzman, 2009). For this reason, CBA IAMs are mostly excluded from the assessment of transformation pathways in AR5 and SR15[3], and are not considered in this thesis.

According to Krey et al. (2019b), large-scale IAMs have become increasingly influential in informing climate policy over the last few years[4]. Beyond IPCC assessments, these IAMs are used to inform decision making, for instance, in policy impact assessment by the European Commission[5] and for national governments in preparing and updating their Nationally Determined Contributions (NDCs). Given the Paris Agreement and the associated process of continuously updating NDCs, it is expected that large-scale IAMs will continue to play an important role in the coming decades (Krey et al., 2019b).

While the role of CBA IAMs in climate policy has been discussed and criticised at length (e.g. Ackerman et al., 2009; Pindyck, 2013; Stern, 2013b; Weitzman, 2009), large-scale IAMs have received relatively less attention. This likely has to do, at least in part, with the complexity and significantly larger size of large-scale IAMs, which can require hundreds or thousands of parametric and structural assumptions (Trutnevyte, 2016). Most generalisations about ensembles of large-scale IAMs are likely to be contradicted by exceptions. Many researchers are also not aware of the distinction between CBA IAMs and large-scale IAMs and cite criticisms of CBA IAMs when discussing large-scale (non-CBA) IAMs without realising that many of these don't apply[6].

Still, large-scale IAMs have received an increasing amount of criticism since AR5 (see Gambhir et al. (2019) for a comprehensive review). Much of this has focused on the widespread use of negative emissions technologies (so-called NETs, also referred to as carbon dioxide removal (CDR) technologies) in transformation pathways generated by large-scale IAMs (Anderson, 2015; Anderson & Peters, 2016; Fuss et al., 2014; Larkin et al., 2018; Smith et al., 2016; P. A. Turner et al., 2018), but there is also an ongoing discussion of whether (Anderson & Jewell, 2019) and how (Grant et al., 2020; Hausfather & Peters, 2020; Mccollum et al., 2020) IAMs should be used to inform climate

---

[3] A few models, such as the WITCH model (Bosetti, Carraro, Galeotti, et al., 2006), has sufficient detail to be included in AR5 and can also be run in CBA mode.

[4] Sometimes these IAMs are also referred to simply as 'climate-policy models' (e.g. Anderson & Jewell, 2019).

[5] https://ec.europa.eu/clima/policies/strategies/analysis/models_en

[6] In particular the criticisms of the damage function and the social discount rate that are used to weigh climate damages against mitigation costs do not apply to large-scale IAMs.

policymaking. Some (e.g. Iyer and Edmonds (2018)), have also pointed out that scenarios are prone to misconceptions because of a lack of guidance from modelers about underlying assumptions and missing information on how to interpret results. Some of the criticism also focuses on the implications of uncertainty in assumptions for the insights that can be drawn from IAMs, which is the focus of this thesis (e.g. Rosen and Guenther (2015) go as far as saying that the level and type of uncertainty of input assumptions render the cost of mitigation unknowable). Older research on and discussions of spurious detail (Funtowicz & Ravetz, 1990; Morgenstern, 1963), the dependence of model outputs on arbitrary assumptions (Keepin & Wynne, 1984), and the 'black box' nature of many large models (Robinson, 1990) are also relevant and still feature in more recent assessments of IAMs (see e.g. Stanton et al. (2008)). Still, little research has been done on IAM ensembles *as ensembles,* and uncertainty assessments based on formal techniques such as the ones used in this thesis are still far in between (Gambhir et al., 2019; Yue et al., 2018). This thesis aims to fill part of this gap[7]. The size and complexity of large-scale IAMs makes studying the robustness and reliability of the insights that can be drawn from them a challenging task. Their increasing influence in global assessments of how to achieve the Paris target, however, also makes this an important task.

AR5 relies heavily on IAMs to generate and assess transformation pathways (IPCC, 2014a). It states the questions to be answered with IAMs as follows:

> "Stabilizing greenhouse gas (GHG) concentrations will require large-scale transformations in human societies, from the way that we produce and consume energy to how we use the land surface. A natural question in this context is what will be the 'transformation pathway' towards stabilization; that is, how do we get from here to there? … The chapter is primarily motivated by three questions. First, what are the near-term and future choices that define transformation pathways, including the goal itself, the emissions pathway to the goal, technologies used for and sectors contributing to mitigation, the nature of international coordination, and mitigation policies? Second, what are the key characteristics of different transformation pathways, including the rates of emissions reductions and deployment of low-carbon energy, the magnitude and timing of aggregate economic costs, and the implications for other policy objectives such as those generally associated with sustainable development? Third, how will actions taken today influence the options that might be available in the future?" (IPCC, 2014a, p. 418).

[7] The rest of this thesis will use 'IAM' to refer to large-scale (non-CBA) IAMs.

In order to compute transformation pathways, IAMs represent key interactions among technologies, human systems, GHG emissions, and the climate in a single integrated framework (IPCC, 2014a). Many assumptions and model choices (discussed in the next two sections) are required to generate transformation pathways using IAMs. This thesis identifies and takes first steps towards assessing challenges associated with drawing robust and reliable insights using IAMs given the uncertainty of these model choices and assumptions. The first part of the thesis (chapters 2-4) focuses on the IAM ensemble used in AR5 and the second part of the thesis (chapters 5-6) investigates predictions based on the Future Technology Transformations (FTT) energy system simulation model.

The first part makes two arguments about IAM ensembles. First, it argues that IAM independence is an important but neglected topic in AR5, which has an important bearing on the robustness of the insights that can be drawn. If the IAMs in the AR5 ensemble are not independent, we cannot know whether agreement in outputs is a sign of robustness or a consequence of shared model choices and assumptions. A method is developed that uses model documentation to construct a model family tree and this method is used to identify likely model dependencies in AR5. Based on this, it is found, many of the IAMs in AR5 are not independent. Second, it argues that the risk of the AR5 IAM ensemble being wrong about the sign of the cost of mitigation is significant. AR5 cost estimates, which are exclusively positive, do not capture the full range of cost estimates observed in the literature, which also includes negative values. A review of the AR5 scenario publications suggests that the reason why none of the IAMs in AR5 predicted net negative costs in AR5 is that few of the mechanisms that typically give rise to net negative costs were included in the AR5 IAMs. It is also shown that the model intercomparison studies that generated most of the AR5 scenarios focused only on aspects that increase the cost of mitigation. Based on all of this, it appears, the AR5 IAM ensemble might be biased against net negative cost results.

The second part of the thesis examines the reliability of the predictions generated by FTT. The FTT energy system simulation model is seen by Mercure et al. (2014; 2016) to offer better predictions of the impacts of policies on the energy system than what energy system optimisation models (ESOMs) do, due in part to the derivation of technology deployment rates based on the theory of technology diffusion. However, even though the assumptions of perfect markets and fully rational agents that underpin the descriptive power of ESOMs have been criticised, and even though technology deployment in ESOMs is determined in part by exogenous constraints, it is not clear that FTT's predictions fare any better. A global sensitivity analysis examining the sensitivity of technology deployment and emissions in the power sector sub-model of FTT (FTT:Power) to investor discount rates, technology lifetimes, technology build times, learning rates, and a model specific scaling factor – a constant representing the time it takes to achieve a full turnover of technologies – shows that FTT:Power predictions are highly sensitive to the scaling factor, whose "true" value is deeply uncertain. This poses an issue for the reliability of best guess FTT:Power predictions.

Based on the importance of model independence for robustness and on the risk of being wrong (associated with both ensemble results and individual model results such as those based on FTT), the thesis concludes that IAM research should aim to incorporate a diversity of model choices and assumptions both in the construction of IAM ensembles and in individual IAMs.

1.1 From temperature targets to transformation pathways

IAMs are used to translate climate policy targets, such as the 2˚C Paris target, into concrete transformation pathways, which depict how different sectors of the economy (e.g. energy, industry, transport, agriculture, forestry) in different regions of the world can change in order to meet the required emissions reduction. Among other things, transformation pathways tell us how, when, and where fossil fuels – which are responsible for the majority of GHG emissions – are replaced by alternative energy sources.

How do IAMs do this? First, temperature targets have to be translated into emissions pathways[8]. This is the domain of climate science, in which uncertainty has led to many different climate models, which yield different temperature projections for the same emissions pathway. By running multiple climate models together to generate 'super-ensembles', which convey the variety of temperature responses produced by different climate models (IPCC, 2014a), however, the probability that a given emissions pathway will lead to warming below a certain level can be computed. This is why temperature targets always come with probabilities; A common interpretation of the Paris climate target, for instance, is a 66% chance of staying below 2˚C.

Climate modelling has shown that the probability of meeting a given temperature target is roughly proportional to cumulative emissions (IPCC, 2014a). This relatively simple relationship has given rise to the popular concept of carbon budgets, which tell us how much more we can emit if we want to achieve a given temperature target (with a certain probability). At the time of AR5, the carbon budget corresponding to a 66% chance of staying below 2˚C was estimated to be between 600 and 1,200 (10-90% range) $GtCO_2$ (Anderson & Peters, 2016)[9].

---

[8] IAMs might also take emissions and GHG concentration levels as inputs.

[9] These estimates have changed since AR5, but the range of uncertainty is still large.

Many emissions pathways, however, are compatible with the same level of cumulative emissions. Emissions could either decrease immediately, allowing for a (relatively speaking) slower rate of emissions reductions, or emissions could continue to increase for a while, thereby necessitating more rapid emissions reductions later on[10]. Even the same emissions pathway, however, can be met using a combination of different mitigation measures. As shown in Figure 1.2, many different gases and sectors contribute to GHG emissions. This means that identical emissions pathways can be achieved by targeting different combinations of gases and sectors. Even if we consider only CO_2 emissions (which account for 76% of GHGs), which eventually have to reach net zero, the timing of emissions reductions in different sectors can still vary. And even within the same sector, mitigation options abound. For instance, within the energy sector (electricity and heat + other energy in Figure 1.2), reductions can stem from changes to energy *demand* or energy *supply*. Even if we narrow it down to emissions reductions in *electricity supply alone*, we still have a choice between multiple low-carbon technologies (such as nuclear, hydro, and renewables) for replacing unabated fossil fuel technologies. In addition to all of this, emissions can also be reduced using NETs. In short, climate mitigation is characterised by a very high degree of freedom and IAMs have to make a multitude of choices in order to go from a temperature target, such as the 2˚C Paris target, to a concrete transformation pathway. These choices will be defined not only by the solution mechanisms that are employed by IAMs, but also the many structural and parametric assumptions that have to be made in order to represent the economic and technological systems that are involved. All of these choices and assumptions involve a great deal of uncertainty.

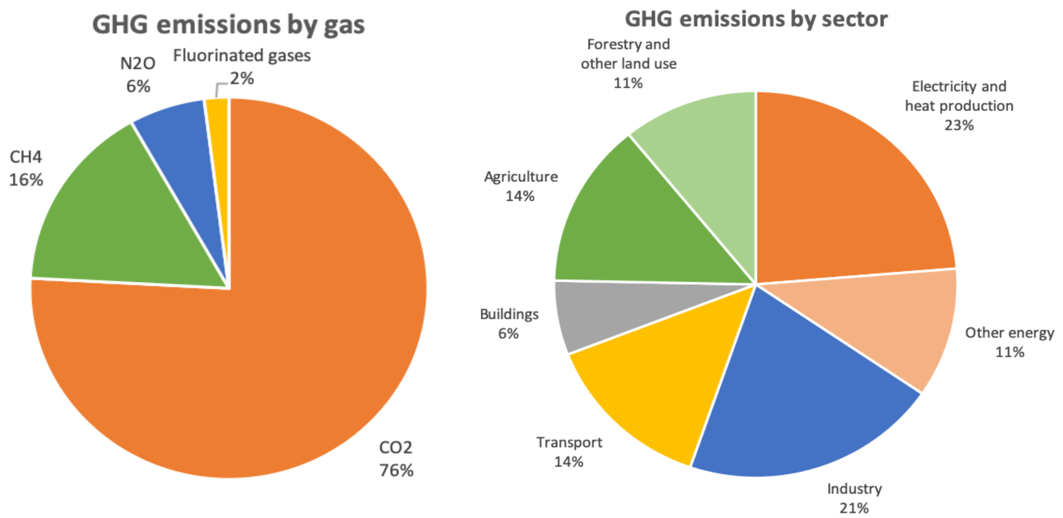


Figure 1.2 Global GHG emissions by gas (left) and by sector (right) in 2010. Numbers are taken from AR5 and show percentages in 2010 (IPCC, 2014a).

[10] Anderson and Peters (2016) warn against relying too heavily on carbon budgets, as relationships between emissions and temperature rises in reality are more complex. It is, however, still true that cumulative emissions provide a good first-order approximation for temperature increases.

1.2 Plausible futures

Any attempt at modelling the economy or the energy system (and the interactions between the two) at time scales relevant to climate change mitigation (typically on the order of 30-100 years) faces a great deal of uncertainty. A footnote in the Summary for Policymakers (SPM) in the AR5 Working Group III (WGIII) report tells us that:

> "[IAMs] are simplified, stylized representations of highly-complex, real-world processes, and the scenarios they produce are based on uncertain projections about key events and drivers over often century-long timescales. Simplifications and differences in assumptions are the reason why output generated from different models, or versions of the same model, can differ, and projections from all models can differ considerably from the reality that unfolds." (IPCC, 2014a, p. 10).

IAMs need to make hundreds, if not thousands, of assumptions regarding model structure and parameter values in order to estimate the value of variables such as the cost of mitigation and the deployment of renewable energy technologies (Trutnevyte, 2016). This introduces two types of uncertainty that will play central roles in this thesis: *parametric uncertainty* and *structural uncertainty*. Parametric uncertainty is the uncertainty associated with the value of model parameters. There are many different sources of parametric uncertainty, such as measurement error and natural variability (for a thorough treatment, see e.g. Morgan and Henrion (1990)). Structural uncertainty is the uncertainty associated with the equations that are used to relate parameters and variables in the model. Structural uncertainty reflects the imperfect and incomplete knowledge of the system being modelled (Boulanger & Bréchet, 2005; DeCarolis, 2011). Both kinds of uncertainty are strongly present in IAMs.

Partly for this reason, AR5 uses the term "plausible descriptions of the future" when talking about IAM scenarios, and points out that "scenarios are neither predictions nor forecasts" (IPCC, 2014a, p. 1270). Instead, IAMs are commonly seen to provide answers to 'what-if' questions (e.g. Anderson & Jewell, 2019; Schneider, 1997), and answers to these questions are often coached in careful language, using phrases such as "what *might* happen" (Mcdowall et al., 2014).

One of the key purposes of asking 'what-if' questions is to gain a better understanding of the consequences of decisions and other developments (Anderson & Jewell, 2019). For this to be the case, however, answers must tell us something about the actual mechanisms involved in climate mitigation. If a 'what-if' question is answered using an inference tool that fails to mimic relevant mechanisms in the real world then it is likely to suggest consequences that are different to those that will actually occur. IAMs only provide information that is useful for decision making if they get some things (and enough of them) right. Thus, even though IAMs do not forecast or predict the future – because the key drivers are too uncertain, and it is impossible to foresee all events that might shape outcomes – we still expect answers to 'what-if' questions to capture real-world relationships to some extent. Otherwise, the answers would simply not be useful. For example, answers to questions about aggregate mitigation costs would not be useful if we had no reason to believe that they capture some of the important cause-effect relationships or at least get the order of magnitude right.

Given the significant uncertainty regarding both parameter values and causal structures, however, different IAMs provide different answers to the same 'what-if' question. This poses challenges for the interpretation of results. While it is clear – and often repeated – that IAMs provide "insights, not numbers" (Peace & Weyant, 2008), there are still important questions to be asked regarding what insights can be gleaned and how. Ultimately, the confidence we place on IAM results matters not only for their usefulness in informing climate policy, but also for framing the public debate. In particular, it matters because the consequences of being wrong, and in that way providing a poor basis for decision making and public deliberation, can be dire.

1.3 Overview of thesis

As noted, IAMs have become central tools in the assessment of global transformation pathways that can meet the Paris target and play a central role in IPCC reports, including in AR4 (IPCC, 2007), AR5 (IPCC, 2014a), and in SR15 (IPCC, 2018a). IAMs will also play a key role in the upcoming IPCC 6th assessment report (AR6)[11]. While the first part of this thesis (chapters 2-4) examines the robustness and uncertainty of IAM ensemble results, the second part of the thesis (chapter 5-6) examines the reliability of individual IAM results.

More specifically, Chapter 2 draws on research conducted in climate modelling (Jun et al., 2008; Knutti et al., 2013; Masson & Knutti, 2011; Tebaldi & Knutti, 2007) to argue that the robustness of insights drawn from IAM ensembles requires independence between IAMs; independence that is assumed without discussion in AR5. The chapter develops a method to construct a model family tree based on similarities between models explicitly stated in their model documentation and uses this method to identify likely dependencies between IAMs in AR5. The analysis shows that the 14 most influential IAMs in AR5 form three branches, the largest of which is the MESSAGE/MERGE branch, which includes MERGE, MESSAGE-MACRO, MERGE-ETL, REMIND, WITCH, and BET. Together, the IAMs in this branch are responsible for about half of the scenarios in the AR5 ensemble. The model documentation is also used to show that IAM development over time has been driven by an expanding

[11] Based on the contents of the first order draft and communication with one of the lead authors, Glen Peters.

set of policy questions that have incentivised a continuous increase in the level of detail and scope of analysis.

Utilising the AR5 WGIII report's discussion of key differences in IAM structure, Chapter 3 shows that there is a strong overlap between the IAM categories that arise when considering the degree of foresight and economic coverage in IAMs, the main model frameworks (optimal growth theory, computable general equilibrium modelling, and energy systems optimization modelling), and the branches in the model family tree constructed in Chapter 2. All the IAMs in the MESSAGE/MERGE branch, for example, are general equilibrium – perfect foresight models based on (Ramsey) optimal growth theory. Chapter 3 also shows, however, that the model family tree captures sources of model dependencies and independencies that the key differences in model structure discussed in AR5 do not capture. This is not surprising given the many sources of model dependencies (some of which are discussed in Chapter 2 for climate models).

Chapter 4 starts from the observation that all the estimates of the cost of mitigation reported in AR5 are positive. Yet, according the literature, the cost of mitigation could also be net negative. Chapter 4 reviews several reasons for how net negative mitigation costs might arise and presents a number of mechanisms that, if included in models, might give rise to net negative cost results. The chapter notes that experts disagree when it comes to the ability of these mechanisms to generate net negative costs in the real world. This means that the AR5 IAM ensemble captures only part of the uncertainty associated with the cost of mitigation. Based on Douglas (2009) and Rudner (1953), the chapter develops an argument for why this uncertainty is important. In short, being wrong about the cost of mitigation could have large negative consequences. This is used to argue that a diversity of model choices and assumptions is important not only for ensuring robustness of insights (as argued in Chapter 2) but for reducing the risk of being wrong.

Chapter 4 also examines potential reasons for why the AR5 IAM ensemble contains only net positive cost results. While some authors have argued that general equilibrium models (including CGE and optimal growth models), which are responsible for all the cost estimates in AR5, exclude net negative costs by construction, Chapter 4 shows that general equilibrium models can be modified in order to reflect the possibility of net negative costs by taking into account the sorts of mechanisms previously identified. A review of the AR5 scenario publications, however, indicates that only two IAMs in AR5 (IMACLIM and, to some extent, WITCH) include any of the mechanisms identified as enabling net negative costs. Additionally, Chapter 4 shows that the model intercomparison studies that are responsible for the vast majority of the scenarios in the AR5 ensemble focused on aspects that can only increase the cost of mitigation. Based on this, Chapter 4 concludes, there is reason to believe that the AR5 IAM ensemble is biased towards net positive mitigation costs.

Chapters 5 and 6 turn to predictions based on FTT, which is seen by Mercure et al. (2014; 2016) to differ from most other energy systems models. Chapter 6 conducts a global sensitivity analysis of the power sector sub-model of FTT (FTT:Power) and Chapter 5 provides the context and motivation for the analysis conducted in Chapter 6. The goal of both chapters is to begin to assess whether FTT meets the goals that it (according to those who built it) was designed to meet.

Chapter 5 presents the claims put forth by FTT modelers regarding the superiority of FTT relative to ESOMs. In particular, the chapter focuses on the claim by Mercure et al. (2014; 2016) that the endogenous derivation of technology deployment rates based on the theory of technology diffusion enables a more realistic depiction of the impacts of policies on future technology deployment. The chapter shows that ESOMs are likely to provide poor predictions if either the assumptions of perfect markets and rationality do not hold (sufficiently), or if the exogenous technology deployment constraints do not accurately capture real-world constraints. However, even though the neoclassical assumptions have been widely criticised and even though technology deployment in ESOMs is determined in part by exogenous constraints, Chapter 5 concludes, we cannot claim that FTT predictions fare any better without (at least) assessing the sensitivity of FTT predictions to key uncertain assumptions.

In order to begin to do so, Chapter 6 conducts a global sensitivity analysis of the FTT:Power model using Monte Carlo analysis and Latin Hypercube sampling. The goal of the chapter is to provide a first conservative estimate of the uncertainty of FTT:Power predictions, and to identify the parameters in the FTT core equation with the largest influence on FTT:Power predictions. The analysis represents the first sensitivity analysis of any of the FTT models that goes beyond a one-factor-at-a-time approach and includes all the parameters that define the core equation (the shares equation). Uniform and independent distributions based on varying default parameter values by ±50% are used to examine the sensitivity of technology deployment and emissions to investor discount rates, technology build times, technology lifetimes, learning rates, and the overall scaling factor – a constant representing the time it takes to achieve a full turnover of technologies. Given that the analysis in Chapter 6 includes only a sub-set of FTT:Power parameters and ignores structural uncertainty (and given that the ±50% parameter ranges are shown to be close to ranges observed in the literature), the results can be interpreted as a first order, conservative estimate of the uncertainty of FTT:Power predictions.

The results show that, for the chosen parameter ranges, technology diffusion alone is not sufficient to reduce emissions in line with the 2˚C pathway in FTT:Power. This suggests (if we accept the structural assumptions in FTT:Power) that policies will be necessary to limit emissions from the power sector. The results, however, also show that the FTT:Power predictions of policy impacts depend crucially on the scaling factor, whose “true” value is deeply uncertain. Thus, the chapter concludes, the influence of the scaling factor on FTT:Power predictions poses an issue for the accuracy of best guess FTT:Power predictions. Given the importance of the uncertainty, Chapter 6 concludes, it is more appropriate to provide policymakers with ranges of results contingent on key parameter values.

Overall, chapters 5 and 6 thus show that while the rates of technology deployment in ESOMs are determined partly by exogenous constraints drawn from limited empirical evidence, the rates of technology deployment in FTT:Power are determined to a large extent by the scaling factor. The value of the scaling factor appears to be no more certain or any easier to verify than the values of the maximum technology deployment rates that are assumed in ESOMs.

Based on the arguments and findings presented in chapters 2-6, Chapter 7 concludes, the IAM community should strive to incorporate a diversity of approaches and assumptions both in IAM ensembles and in individual IAMs. This is key to ensuring robustness of insights and reflecting important uncertainties associated with IAM research.

2 IAM dependencies in AR5

A key aim of the IPCC reports is to gather and assess the knowledge base that exists in the areas of climate change (Working Group I), mitigation (Working Group III), and adaptation (Working Group II). Representing the uncertainty of the knowledge base is central to this. According to the AR5 Synthesis Report,

> “[a]n integral feature of IPCC reports is the communication of the strength of and uncertainties in scientific understanding underlying assessment findings. Uncertainty can result from a wide range of sources… Complex interactions among multiple climatic and non-climatic influences changing over time lead to persistent uncertainties, which in turn lead to the possibility of surprises.” (IPCC, 2014b, p. 37)

As mentioned in Chapter 1, different climate models predict different temperature responses for the same emissions pathway. The variety of responses is a result of the complexity of the climate system and associated uncertainties regarding how to model it (Tebaldi & Knutti, 2007). For this reason, the IPCC uses *ensembles* (sometimes called ‘super-ensembles’) of climate models to capture the differences in temperature responses predicted by different climate models. The use of climate model ensembles to estimate the likely impacts of emissions on temperature increase is seen as a way of capturing the structural uncertainty associated with modelling climate change.

Similarly, the IPCC uses *IAM ensembles* to assess transformation pathways, i.e. pathways depicting how we might reduce GHG emissions in order to meet different climate targets. Just like ensembles of climate models can be used to show the spread of temperature responses generated by different climate models (and different versions of the same climate models), ensembles of IAMs can be used to show the spread of different IAM outputs generated by different IAMs (and different versions of the same IAMs).

A key question is whether the spread in outputs captured by IAM ensembles provides a good representation of the structural uncertainty that is associated with integrated assessment modelling. Based in part on a series of papers published in the climate modelling literature, this chapter argues that IAM independence is crucial for our ability to draw robust insights from IAM ensembles. In order to assess model dependencies between the IAMs in the AR5 ensemble, this chapter develops a method that deliberately avoids detailed bottom-up comparisons of IAMs, because the number and complexity of IAMs in AR5 would render this infeasible. The method – which is based on document analysis – is used to construct a family tree that indicates *likely model dependencies* in AR5. In addition to this, the documents are used to obtain a richer understanding of what has driven the development of IAMs over the last few decades.

The results of the analysis indicate that a significant number of model dependencies exist within the AR5 IAM ensemble. Given the importance of the independence assumptions for our ability to obtain robust insights, this chapter argues, future IAM research and IPCC reports should pay much more attention to IAM dependencies.

Section 2.1 introduces the AR5 IAM ensemble. Section 2.2 explains how model independence represents an important, but rarely justified, assumption in climate model ensembles. Section 2.3 shows that the AR5 WGIII report mentions dependencies between scenarios that stem from the same IAM and the same model intercomparison study but says nothing about dependencies between IAMs. This, section 2.4 argues, presents an issue for the interpretation of agreement among IAMs in an ensemble as a sign of robustness. Section 2.5 presents the method developed in this chapter to identify likely model dependencies in AR5 and section 2.6 presents the resulting model family tree. Section 2.7 presents the results of the additional document analysis, which indicates that the evolution of IAMs has been driven by policy demands that appear to have incentivised an increase in detail and scope, rather than a diversity of modelling approaches. Section 2.8 provides concluding remarks.

2.1 Using IAM ensembles to capture structural uncertainty

Figure 2.1 provides an example of how scenarios produced by the IAMs in the AR5 ensemble are presented in AR5. The top panel shows global GHG emissions over time in all the scenarios in the ensemble, and the bottom panel shows the shares of low-carbon energy in scenarios grouped according to the GHG concentration levels achieved. The bottom right figure, for example, shows the share of low-carbon energy in all the scenarios that reach GHG concentration levels of 430-480 ppm CO_2eq. Not surprisingly, because this group includes the scenarios that achieve the lowest GHG concentration levels, the average low-carbon shares (in each year) in this group is higher than the average low-carbon shares (in the corresponding years) in the other groups. The ranges, however, are still large. This illustrates how scenarios can achieve the same GHG concentration levels with different shares of low-carbon energy.

The ranges above are partly a consequence of the many degrees of freedom in IAMs, which means that the same climate target can be met in many ways. It is also, however, a result of the uncertainties associated with modelling technological and human systems, such as the economy and the energy system. The value of an IAM output is determined by the value of IAM inputs (e.g. assumed population growth, policies), the value of IAM parameters (e.g. technology costs, rates of renewable technology cost reductions), and model structure (e.g. whether renewable technology cost reductions are exogenous or endogenous, how decision making is modelled). If either input values, parameter values, or model structure is changed, output values will change too[12]. Due to our incomplete knowledge regarding how to model technological and human systems, we don't know the "most correct" or "best" way to do so. Because different IAMs encompass different assumptions and modelling choices (materialising in different model structures and parameters), different IAMs will generally compute different output values for the same input values. IAM ensembles are thus seen to capture some of the inherent uncertainties involved in integrated assessment modelling. Capturing *structural uncertainty* is one of the main reasons for including many *different* IAMs in the same ensemble.

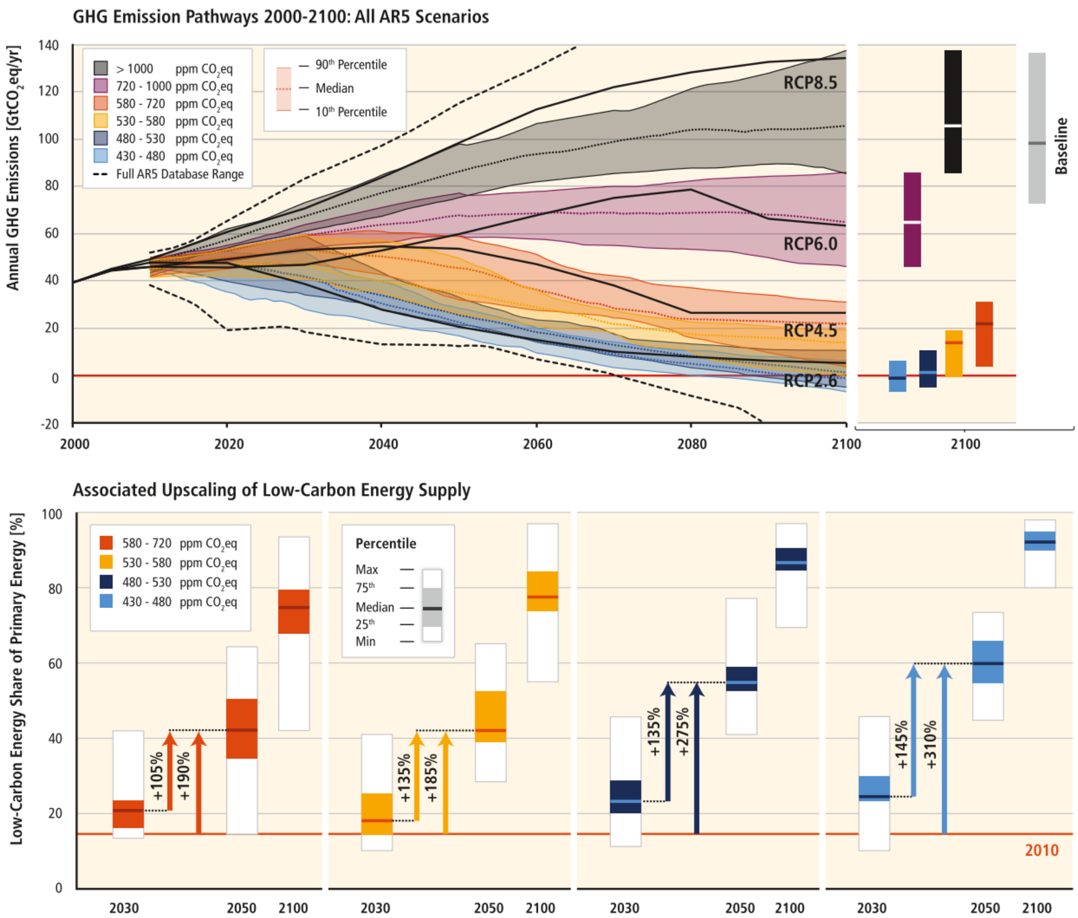


Figure 2.1 Global GHG emissions in all scenarios in the AR5 ensemble (upper panel) and associated upscaling requirements of low-carbon energy for 2030, 2050, and 2100 compared to 2010 in scenarios that achieve different GHG concentration levels (lower panel). The upper panel also shows the Representative Concentration Pathways (RCPs), which correspond to different levels of radiative forcing (measured in W/m^2) in 2100. GHG concentration levels are indicated by the legends. Reproduced from IPCC (2014a, p. 11).

[12] It is possible, in principle, that changing two or more of these aspects simultaneously will generate the same output value.

Scenarios have been used by the IPCC since the first assessment report (IPCC, 1990). At that point, only two different IAMs were used to generate emissions scenarios. Since then, the number of IAMs that generate scenarios have increased significantly. In AR5, "the number of models has increased, and model functionality has significantly improved since AR4, allowing for a broader set of scenarios in the AR5 ensemble" (IPCC, 2014a, p. 423). All scenarios included in the IPCC reports have to be published in the peer-reviewed literature. Based on this literature, a total of 1,184 scenarios from 30 different IAMs were collected for the AR5 scenario ensemble[13]. The majority (about 95%) of these scenarios were generated as part of nine model intercomparison studies (IPCC, 2014a). Model intercomparison studies are studies involving many different models and modelling teams aimed at investigating specific topics. The two model intercomparison studies responsible for the largest number of scenarios in the AR5 Scenario Database, AMPERE (Kriegler et al., 2015b; Riahi et al., 2015a) (responsible for 378 scenarios) and EMF 27 (G. J. Blanford et al., 2014; Krey et al., 2014; Kriegler et al., 2014) (responsible for 362 scenarios) explore the implications of delayed and fragmented mitigation and of limitations on technology (including on costs, performance, and availability). RoSE (Bauer et al., 2016; Calvin et al., 2016; Chen et al., 2016; De Cian et al., 2016; Luderer et al., 2016) (responsible for 105 scenarios), LIMITS (Kriegler et al., 2013; Tavoni et al., 2014) (responsible for 84 scenarios), and EMF 22 (Clarke et al., 2009) (responsible for 70 scenarios) explore delayed and fragmented mitigation, and ADAM (Edenhofer, Knopf, Leimbach, & Bauer, 2010) (responsible for 15 scenarios) explores the implications of limitations on technology. Lastly, RECIPE (Luderer et al., 2012) (responsible for 18 scenarios) again explore the implications of delayed and fragmented mitigation and limitations on technology. A key question is whether a continuously increasing number of IAMs (and, related, an increasing number of scenarios) leads to better insights.

This chapter argues that dependencies between IAMs in the AR5 ensemble is an important but neglected topic. It is important because it influences the robustness of insights that can be gleaned from scenario ensembles. Such dependencies have already been shown to be prevalent in climate model ensembles, which in many ways have provided a template for IAM ensembles (Knutti et al., 2013; Masson & Knutti, 2011).

2.2 Lack of independence in climate model ensembles

The main motivation for using ensembles of climate models is to capture uncertainty related to the choice of model design (Tebaldi & Knutti, 2007), that is, *structural uncertainty*. In several areas of research, such as public health and agriculture, forecasts based on combining different models have been shown to outperform single-model forecasts (Tebaldi & Knutti, 2007). Similarly, in climate

[13] Available online at https://tntcat.iiasa.ac.at/AR5DB/dsd?Action=htmlpage&page=about

modelling, averages of results from multiple models have been found to agree better with observations than results from single models.

The reason for this is essentially this: in an ensemble, errors introduced via different modelling choices and assumptions tend to cancel out. Errors in modelling are an unavoidable result of the fact that models are simplifications of the real world. They reflect the fact that correspondences between the models and the real world are never one-to-one. Specifically, “simplifications, assumptions and choices of parametrizations have to be made when constructing a model, and they inevitably lead to errors in the model and the forecasts it produces” (Tebaldi & Knutti, 2007, p. 2056). By combining different models, the errors of different models can be made less severe. Based on a reading of the law of large numbers, if models are assumed to be normally distributed around an error free mean, then as more and more outputs from different models are aggregated, the errors of different models will cancel out.

This, however, rests on a crucial assumption, namely that model choices and assumptions in the different models are made *independently of each other*. Without independence, we cannot assume that errors are normally distributed around an error free mean. Quite the contrary, if model choices and assumptions are dependent, errors will be correlated. In this case, errors will not cancel out. Put differently, the models’ “deviations from the true system or other models will be similar” (Masson & Knutti, 2011, p. 1).

A handful of papers have presented methods to show that many climate models used in climate model ensembles are not independent. By comparing the outputs of 20 state of the art climate models with observations, Jun et al. (2008) provide evidence that many climate models have highly correlated errors. Knutti and Masson (2011) develop a distance metric based on differences between modelled and observed temperature and precipitation levels in order to measure climate model dependence. Both papers find strong similarities between many models.

Overall, Knutti and Masson (2011) find, dependencies exist between models developed at the same institution, between models sharing similar components, and between successive versions of the same model. This is also consistent with Jun et al.’s (2008) finding that climate models developed by the same institution have similar errors. In many ways, this is not surprising. New climate models are rarely written from scratch but “evolve from combining, modifying and improving existing parts and ideas” (Masson & Knutti, 2011, p. 2). Some institutions use entire model components from other models. What this means is that the effective number of independent models in climate model ensembles is lower than the actual number of models.

Based on their findings, Knutti et al. (2013) conclude that the independence assumption in climate model ensembles is rarely justified. Rather, "current coordinated model experiments are like asking the same question to a small number of people, without thinking about how to select those people, how many to ask, and how to account for the fact that they may have similarly biased opinions. This undoubtedly makes the interpretation of the answers challenging" (Masson & Knutti, 2011, p. 4).

2.3 Dependencies in the AR5 ensemble

Chapter 6 in the AR5 WGIII report acknowledges that ensemble results are not straightforward to interpret. In particular, it notes the "unavoidable ambiguity in interpreting ensemble results in the context of uncertainty" (IPCC, 2014a, p. 423). This ambiguity arises in part because "the scenarios assessed in this chapter do not represent a random sample that can be used for formal uncertainty analysis" (IPCC, 2014a, p. 423). The reason for this is twofold.

First, the vast majority of the scenarios in the AR5 ensemble were generated as part of model intercomparison studies. As already mentioned, these studies are focused on exploring particular questions such as the consequences of delayed climate action or limited technology availability. They tend to impose specific assumptions and often harmonise key parameters in order to make results comparable. In addition to this, several scenarios also represent sensitivity runs (for instance with respect to different levels of technology availability). Because "each scenario [in the AR5 ensemble] was developed for a specific purpose" it follows that "the collection of scenarios…does not necessarily comprise a set of 'best guesses'" (IPCC, 2014a, p. 423). This means that the spread of the scenarios cannot be taken to represent the uncertainty of the most likely outcome.

Second, some modelling groups have generated significantly more scenarios than others. Since each scenario is weighted equally in AR5, IAMs that contribute with more scenarios have a larger influence on ensemble results (such as averages). According to Chapter 6, this introduces "a weighting of scenarios that can be difficult to interpret" (IPCC, 2014a, p. 423).

Using the language of the previous sections, the model choices and assumptions made in scenarios from the same model intercomparison study or from the same IAM are not independent. In other words, these scenarios are not independent. In this sense, AR5 acknowledges some of the dependencies that complicate the interpretation of ensemble results.

AR5 does not, however, say anything about possible dependencies *between IAMs*, and thus about scenarios that stem from dependent IAMs (see Table 2.1). If anything, AR5 highlights the independence of different modelling groups (IPCC, 2014a, p. 175). Several of the model intercomparison studies that contribute to the AR5 ensemble also highlight the diversity of the IAMs included (Kriegler et al., 2013, 2015b; Riahi et al., 2015a). For instance, in presenting the AMPERE model intercomparison study (which is responsible for 32% of the scenarios in AR5), Riahi et al. (2015b, p. 12) point out that "the diversity of approaches is an important asset, since it helps us to better understand structural uncertainties, and to focus on findings that are robust across a wide range of methodologies."

Table 2.1 Three ways in which scenarios might be dependent.

Scenarios from…	**Are according to AR5…**
The same IAM	Not independent
The same model intercomparison study	Not independent
Dependent IAMs	[Not mentioned]*

* The work in this chapter aims to fill this gap.

In the end, the AR5 WGIII report "does not attempt to resolve the ambiguity associated with ranges of scenarios" but instead "focuses simply on articulating the most robust and valuable insights that can be extracted given this ambiguity" (2014a, p. 424). As we will see in the next section, however, the concept of robustness again relies crucially on independence.

2.4 Agreement as robustness

A robust model result is a result that is invariant with respect to uncertain model assumptions, be they parametric or structural. If model A, based on assumptions $\{a_1,\ldots,a_n\}$, produce the same value of an output as model B, which is based on equally plausible assumptions $\{b_1,\ldots,b_n\}$, then the value of the output can be said to be robust with respect to variations in these assumptions. The concept is most straightforward when applied to parametric assumptions. If we are faced with a range of plausible parameter values and incomplete knowledge with respect to choosing a value, it is good practice to check whether model results depend on the choice. If results depend (significantly) on the choice, the results are said to be sensitive to the parameter in question. If results don't depend (significantly) on the choice, the results can be said to be robust with respect to the parameter in question. In order to establish robustness, however, we have to make sure that the results are (sufficiently) invariant with respect to the entire range of plausible parameter values. If we only check what the results are for a smaller section of the plausible parameter range, we cannot conclude that results are robust with respect to the uncertainty of the parameter in question. Although the concept of robustness gets more complicated when applied to structural assumptions, a similar point can be made: we cannot conclude that results are robust with respect to structural assumptions unless we make sure that results are (sufficiently) invariant with respect to all plausible structural assumptions.

In theory, then, the requirement for defining insights as robust is that they are (sufficiently) invariant to *all plausible modelling choices and assumptions*[14]. In practice, this requirement is of course not possible to attain for IAMs: We can never be sure that all possible modelling choices and assumptions are captured in IAM ensembles. The point remains, however, that the ability to obtain robust insights from IAM ensembles relies on a *sufficiently diverse* ensemble. If too many IAMs are dependent, diversity will be curbed.

Providing robust insights is an important aim of IAM research and a key justification for including many IAMs in ensembles. This is highlighted both in AR5 itself and in model intercomparison studies on which AR5 relies. Moreover, results from the AR5 ensemble are communicated to policymakers and the wider public in the Summary for Policy Makers (SPM) and the synthesis report without the caveats listed in Chapter 6 of the WGIII report. When IAMs agree on results, this is likely to be perceived by the public as a sign of high confidence in those results. But this interpretation of agreement in results is, as we have seen, questionable if many IAMs are dependent. In this case, agreement in results might be a consequence of shared modelling choices and assumptions rather than a consequence of robustness with respect to a variety of modelling choices and assumptions.

The concepts of *agreement* and *independent lines of evidence* are, in fact, already key to the assessment of uncertainty in the IPCC:

> "The IPCC Guidance Note on Uncertainty defines a common approach to evaluating and communicating the degree of certainty in findings of the assessment process. Each finding is grounded in an evaluation of underlying evidence and agreement. In many cases, a synthesis of evidence and agreement supports an assignment of confidence, especially for findings with *stronger agreement and multiple independent lines of evidence*. The degree of certainty in each key finding of the assessment is based on the type, amount, quality and consistency of evidence (e.g., data, mechanistic understanding, theory, models, expert judgment) and the degree of agreement" (IPCC, 2014b, p. 37 my italics).

In other words, agreement on its own is not enough to assign a high degree of confidence to a finding. The agreement must be based on independent lines of evidence. If scenarios are added to IPCC

[14] Being invariant to *implausible* modelling choices and assumptions is not a requirement. If the value of a parameter is well known – that is, if the uncertainty is insignificant – the only plausible value of the parameter is the given value. Thus, insights can be robust even though they are sensitive to an assumption as long as there is little or no uncertainty regarding that assumptions.

ensembles from IAMs that are not independent, results (such as those shown in figure 2.1) might appear less spread out, and thus less uncertain. An increase in the number of dependent IAMs might thus lead to a false impression of increasing confidence. In reality, if bias is defined as systematic error in a specific direction, ensembles that contain many dependent models will be biased in the direction of the errors associated with the shared modelling choices and assumptions that are responsible for those dependencies. Thus, the degree of independence between the IAMs in the AR5 ensemble should be considered when evaluating the quality of the evidence referred to in the above quote.

2.4.1 Ensembles of opportunity

Since only scenarios published in the peer-reviewed literature can be included in IPCC reports, IAM ensembles, like climate model ensembles, represent "ensembles of opportunity in which the sampling and dependence in the model space is unknown" (Masson & Knutti, 2011, p. 1). The use of ensembles represents a pragmatic approach that allows the IPCC to capture some of the structural uncertainty that is associated with integrated assessment modelling. The ensemble approach has arisen in climate modelling due in part to a lack of unique model quality metrics (Masson & Knutti, 2011). If we knew what models were "more correct" or "better" (in a given context), then we would simply use these. Such quality metrics are, if anything, even harder to come upon for IAMs. In general, IAMs resist the kind of validation that models of physical processes can be exposed to (DeCarolis, 2011)[15]. Among other things, whereas climate model outputs are regularly compared to observations, IAM outputs are not. This makes it even more difficult to determine the quality of IAMs (Decarolis et al., 2012).

Thus, it makes sense to combine the results of many different IAMs and to try to draw insights based on this. But even though IPCC reports rely on results that have been published in the peer-reviewed literature, they have a mandate to assess the strength and uncertainties of those results. Given how model dependencies impact on the robustness of insights that can be obtained from IAM ensembles, IPCC reports should pay more attention to these.

The issues posed by IAM dependencies and the absence of an assessment of such dependencies highlights a gap, namely the need for methods to identify such dependencies. The next section presents the new method developed in this chapter to identify likely model dependencies in IAM ensembles.

[15] Hodges and Dewar (1992) describe the problem of model validation and explains why some models cannot be validated.

2.5 IAMs in AR5

Figure 2.2 shows the number of scenarios generated by each of the 30 IAM in the AR5 ensemble based on data provided in Appendix II.10 of the AR5 WGIII report. As already mentioned, the 30 IAMs vary in terms of how many scenarios they contribute with. Because each of the 1,184 scenarios is given equal weight in the AR5 ensemble, this means that some IAMs have a larger influence on ensemble results than others. Figure 2.2 shows that the most influential IAM in AR5[16], with 158 scenarios (13% of the AR5 ensemble), is the REMIND model. On the opposite end, the Ecofys Energy Model contributes with only one scenario (less than 0.1% of the AR5 ensemble). In particular four models stand out by contributing with more than 100 scenarios each: REMIND, MESSAGE, GCAM, and WITCH. Together, these four IAMS – from now on referred to as the 'Big Four' – are responsible for almost half of the scenarios in the AR5 ensemble.

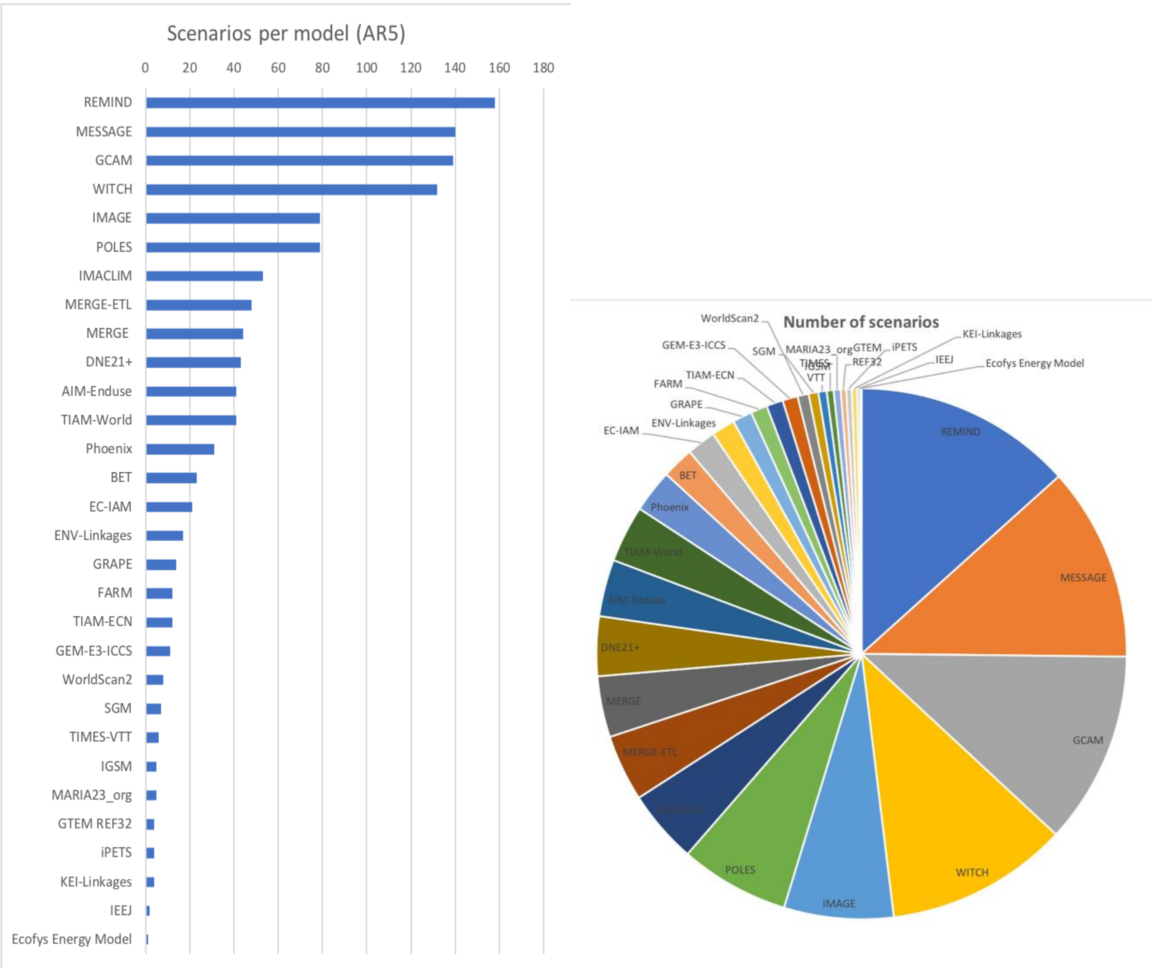


Figure 2.2 Number of scenarios generated by each IAM in the AR5 scenario ensemble.

[16] When influence is measured as the number of scenarios a model contributes with.

Scenarios that stem from the same IAM are not independent (given that they, with the exception of smaller variations, are based on the same model choices and assumptions). Figure 2.2 thus shows that relatively large groups of scenarios in AR5 are not independent. Correlations between scenarios from certain models, especially from the Big Four, might therefore be important to consider when interpreting AR5 ensemble results. Despite acknowledging the ambiguity that an uneven number of scenarios introduces, AR5 does not show how uneven this distribution is, nor does it say anything about what IAMs are most influential.

There are, nonetheless, still 30 different IAMs contributing with scenarios in AR5. If these were all independent, the uneven distribution of scenarios per model might not pose such a big problem (13.5% is, after all, still a relatively small fraction of all the scenarios in the ensemble). An even more important question, therefore, is the extent to which scenarios from different IAMs may also be correlated due to dependencies between IAMs.

Before describing the method used to identify likely model dependencies, it should be noted that all the IAMs in the AR5 ensemble are large and complex. They typically involve hundreds or thousands of assumptions and parameters and it can take years to achieve a good understanding of even a single of these IAMs. These aspects might deter researchers from attempting to study entire IAM ensembles. It is, nonetheless, no less important to do so. But choices must be made in order to make the task feasible.

When studying an entire ensemble of IAMs, it is not feasible to conduct a detailed bottom-up investigation of all the model choices and assumptions that are made in all of the IAMs. A different, and in some ways more "superficial", approach must be taken. The method developed in this chapter is based on using model documentation to draw up a "model family tree" and to gain a qualitative understanding of how IAMs have evolved. In so doing, the uneven distribution of scenarios per IAM in AR5 was utilized to limit the number of IAMs while still making sure that the vast majority of the scenarios in the AR5 ensemble was captured. The data behind figure 2.2 shows that the bottom 16 IAMs are responsible for only 10% of the scenarios in the AR5 ensemble. Thus, by selecting only the 14 most influential IAMs, it was possible to limit the number of IAMs considerably but still capture 90% (1,051) of the scenarios in the AR5 ensemble.

Limiting the number of IAMs allowed for a more detailed analysis overall. For each of the 14 IAMs, model documentation going back all the way to the very beginning was gathered. First, references to the journal papers that present the scenarios in AR5 were obtained from the WGIII AR5 Scenario Database (IAMC, 2014). Second, references to model documentation (journal papers, model manuals, working papers, websites) that describe the IAMs used to generate the scenarios were obtained from the journal papers. Based on this, references to earlier model documentation was found. This snowballing approach was employed for each of the 14 IAMs until no earlier documentation could be found. Based on the model documentation gathered through this process, a model family tree providing a “genealogy” of the 14 IAMs was constructed, and a qualitative understanding of how the 14 IAMs have evolved was obtained. Details of the 14 IAMs, including the earliest documentation, are shown in Table 2.2. The results of the analyses are presented in sections 2.6 and 2.7.

Table 2.2 Details of the 14 IAMs responsible for the largest number of scenarios in AR5. Sources: AR5 Appendix II.10 (IPCC, 2014a) and individual model documentation.

Model	**# Scenarios**	**Full name**	**Institution**	**Earliest Documentation**
REMIND	158	Regionalized Model of Investments and Development	Potsdam Institute for Climate Impact Research (PIK) (Germany)	(Leimbach et al., 2010)
MESSAGE-MACRO*	140	Model for Energy Supply Strategy Alternatives and their General Environmental Impacts	International Institute for Applied Systems Analysis (IIASA) (Austria)	(Messner & Schrattenholzer, 2000) (for MESSAGE, the earliest documentation is (Agnew et al., 1979a))
GCAM	139	Global Change Assessment Model	Pacific Northwest National Laboratory (PNNL) (US)	(Edmonds & Reilly, 1983a)
WITCH	132	World Induced Technical Hybrid	Fondazione Eni Enrico Mattei (FEEM) (Italy)	(Bosetti, Carraro, Galeotti, et al., 2006)
IMAGE	79	The Integrated Model to Assess the Global Environment (formerly: Integrated Model for the Assessment of the Greenhouse Effect)	Netherlands Environmental Assessment Agency (PBL) (Netherlands)	(Rotmans, 1990)
POLES	79	Prospective Outlook on Long-term Energy Systems	European Commission’s Join Research Centre (JRC) (Europe)	(Lesourd et al., 1996)
IMACLIM	53		Center for International Research on Environment and Development (CIRED) (France)	(Baron & Salles, 1991)
MERGE-ETL	48	MERGE- Endogenous Technical Learning		(Kypreos & Bahn, 2003)
MERGE	44	Model for Evaluating Regional and Global	Stanford University (US)	(Manne & Richels, 1992)

		Effects of GHG reduction policies		
DNE21+	43	Dynamic New Earth 21+	The Research Institute of Innovative Technology for the Earth (RITE) (Japan)	(Fujii & Yamaji, 1998)
AIM-Enduse	41	Asia-Pacific Integrated Model - Enduse	National Institute for Environmental Studies (Japan)	(Kainuma et al., 1995)
TIAM-World	41	TIMES Integrated Assessment Model - World	International Energy Agency (IEA)	(Richard Loulou & Labriet, 2008)
Phoenix	31		Pacific Northwest National Laboratory (PNNL) (US)	(Brenkert et al., 2004)
BET	23	Basic Energy systems, Economy, Environment, and End-use Technology Model	Central Research Institute of Electric Power Industry (Japan)	(Yamamoto et al., 2014)

*Although MESSAGE-MACRO is referred to simply as MESSAGE in AR5, the full name is used in this thesis to distinguish this IAM from the energy system sub-model MESSAGE, which is depicted separately in the model family tree.

2.6 The model family tree

Figure 2.3. shows the model family tree that was constructed based on the model documentation. The tree is based on information about i) the year in which each IAM was constructed (taken as the year in which the earliest model documentation appears, unless otherwise specified in the documentation), ii) changes to model names over time, iii) precursor models, and iv) links between models (discussed in more detail in section 2.6.1). Whenever the model documentation collected through the snowballing approach gave inconclusive answers, additional model documentation was sought. In total, 117 different sources of documentation for the 14 IAMs and their precursor models were gathered to construct the model family tree.

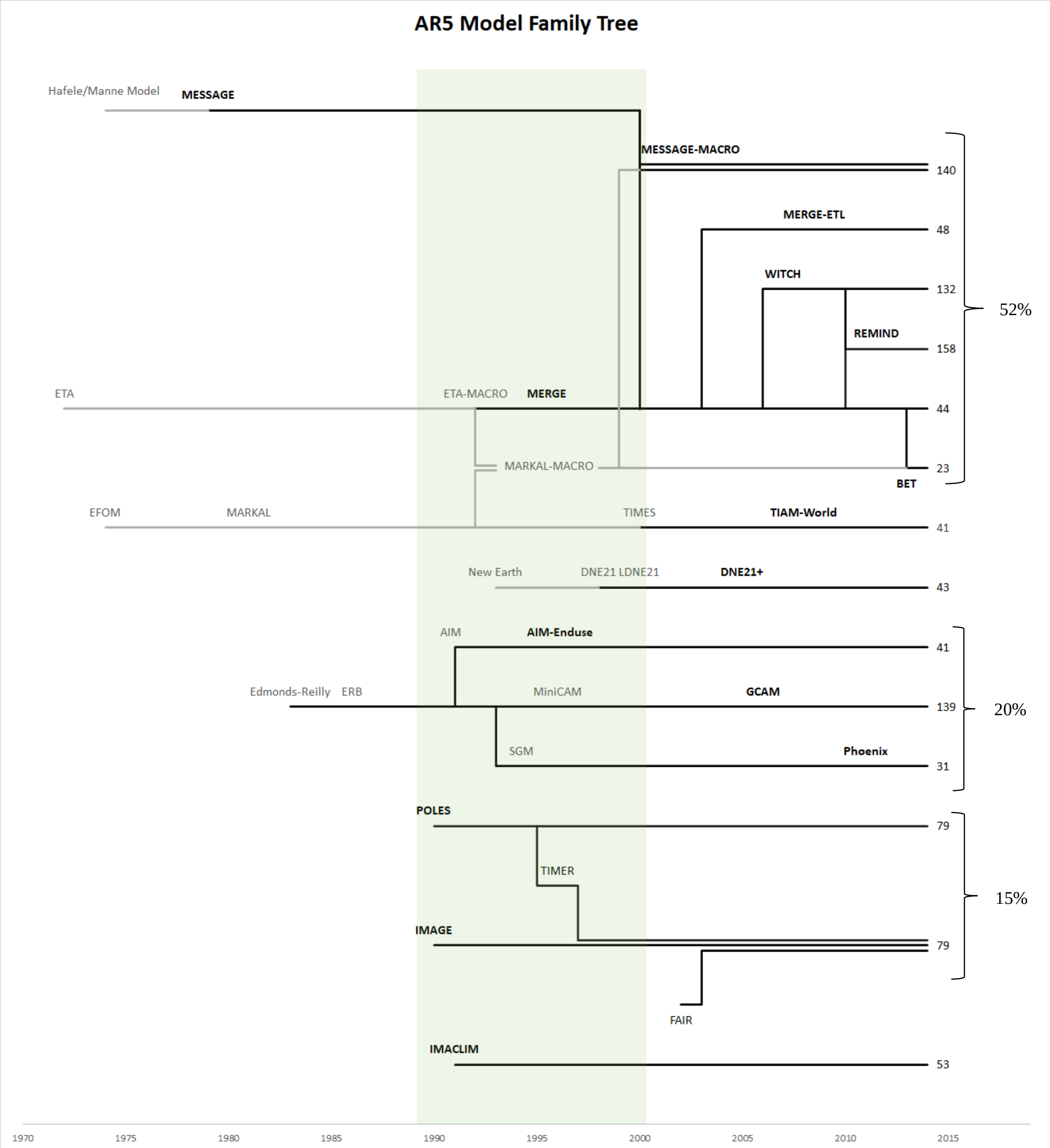


Figure 2.3 Model family tree for the 14 most influential IAMs in AR5 (black lines) including precursor models (grey lines) showing when each model was constructed (the starting point of each line). Vertical lines represent links from newer to older models. These are always placed at the year in which the newer model is constructed. The numbers at the end of each line show the number of scenarios generated by the respective IAMs in the AR5 ensemble. The percentages show the fraction of scenarios generated by the respective branches as a proportion of the total number of scenarios (1,051) generated by the 14 IAMs. The green rectangle highlights the most active decade (1990-2000) in terms of new IAM development.

In short, the model family tree shows how the 14 IAMs that are responsible for 90% of the scenarios in the AR5 ensemble have evolved as a group of models. More specifically, each line in figure 2.3 represents the evolution of one IAM (or, in the case of TIMER and FAIR, sub-models that form components of IAMs). The starting point of each line indicates the year in which the model was constructed. Unless specified in the model documentation, this is taken as the year in which the earliest model publication appears. All the lines end at the year of the AR5 publication. Double and triple lines are used for IMAGE and MESSAGE-MACRO to represent the evolution of sub-models. Model names in black correspond to the names used in AR5, and names in grey show previously used model names. Each name is positioned at the year in which it first appears in the model documentation. Precursor models – models that are intimately related to or considered foundational to but distinct from the AR5 models – are represented using grey lines. The Hafele/Manne model, for example, is a precursor to MESSAGE because it is referred to as "the grandfather of MESSAGE" (Agnew et al., 1979b, p. 4).

Finally, the vertical lines in figure 2.3 represent model links (discussed in more detail in the next sub section). Each link is based on one or several statements in the model documentation that connect the newer IAM to the older IAM. Thus, the source of a given model link is always the model documentation for the newer of the two models that make up a link. For instance, Bosetti et al. (2006), connect WITCH to MERGE when presenting WITCH for the first time. Only links between the 14 IAMs (and precursor models) are included in figure 2.3 (i.e. links to IAMs other than the 14 IAMs are not included). This is because the focus here is on likely dependencies between the IAMs in the AR5 ensemble.

The model family tree shows at least two things. First, it shows when the 14 most influential IAMs in AR5 were constructed, including their precursor models, and as such provides a picture of the evolution of these IAMs as a group. There is a large spread in the age of the 14 IAMs, with the oldest (MESSAGE) dating back to 1979 and the newest (BET) appearing for the first time in 2014, the same year in which AR5 was published. (In terms of precursor models, the oldest model (ETA) dates back to 1972.) The majority of the 14 IAMs were, however, constructed in the 1990s. A likely explanation for this is the publication of the first IPCC assessment report in 1990 (IPCC, 1990) and the supplementary IPCC report that included IAMs for the first time two years later (IPCC, 1992).

Second, the model family tree shows links between the 14 IAMs and, importantly, how these links lead to distinct branches. The largest branch is the branch that stem from the MERGE model, which includes MESSAGE-MACRO[17], MERGE-ETL, WITCH, REMIND, and BET. Together, the IAMs in this branch, which we can call the 'MESSAGE/MERGE branch' are responsible for 52% of the 1,051 scenarios generated by the 14 IAMs. In addition to this, the forerunner to TIAM-World, MARKAL, and MERGE are also linked to BET and MESSAGE-MACRO via MARKAL-MACRO. If we include TIAM-World in the MESSAGE/MERGE branch, the fraction of scenarios in the AR5 ensemble that

[17] MESSAGE-MACRO also links back to MESSAGE (and before this, the Hafele/Manne model).

stem from this branch increases to 56%. Interestingly, all the IAMs that have been constructed after 2000 belong to this branch.

The second largest branch is the branch that stems from the ERB model, which includes GCAM, AIM-Enduse, and Phoenix. The number of scenarios generated by IAMs in this branch amounts to 20% of the scenarios generated by the 14 IAMs. The third largest branch is made up of POLES and IMAGE (connected via TIMER, which is a sub-model of IMAGE). This branch is responsible for 15% of the scenarios generated by the 14 IAMs. Lastly, two models, DNE21+ and IMACLIM, were not found to be linked to any of the other 14 IAMs.

2.6.1 The model links

As noted, the model links in figure 2.3 are based on statements in the model documentation. It is important to understand that these links capture a variety of relationships. What these links have in common, however, is that they capture the *modelers' judgments* – as represented in the model documentation (which is written by modelers) – regarding how their IAM relates to other IAMs at the time of writing. To provide more information about what the links represent, Table 2.3 provides a typology that has been developed inductively based on the statements on which the model links are based. Representative quotes are also shown. Although the quotes in Table 2.3 are almost all (except for AIM-Enduse) taken from the earliest model documentation, the same or very similar statements are often also found in later model documentation (indicating that relationships don't change much and/or that model descriptions are often reused). Additional statements and information related to the links between the 14 IAMs can be found in Appendix A.

'Combination' is used to denote links that are formed two existing models are combined to create a new model. MESSAGE-MACRO, for example, was constructed by combining the energy system model, MESSAGE, with the macroeconomic model, MACRO (which represents a core component of the MERGE model). 'Version' is used when the model documentation indicates that the model in question can be considered a version of another model. MERGE-ETL, for example, is described as "a MERGE model with endogenous technological progress" and Phoenix is described as a "process-level version of the ERB" in the model documentation. For BET, the documentation provides multiple statements that together suggest that BET can be seen as a version of MERGE: BET is described as "a MERGE with advanced, electric end-use technologies" and is developed based on the computer code for MERGE. The label 'Structural' is attached to a link when the model documentation indicates similarities in structure *and* when the stated differences refer only to aspects such as model resolution or the way in which sub-model are combined. REMIND, for example, shares the same intertemporal structure as MERGE but offers a higher degree of technological detail. 'Component' is used when a model uses a component from another model. The only example of this is AIM-Enduse, which has a component that is based on the ERB model. Lastly, 'Similar' is used when the model documentation states that the new model is "similar to" or "has much in common with" another model, but without saying much more about how or in what ways. Since the link types are not mutually exclusive, links can be given more than one label.

Table 2.3 Overview of the model links shown in the model family tree (figure 2.3) including link types and example statements from the model documentation.

Link to	**Link from**	**Type of link**	**Representative Quote**
WITCH	MERGE	Structural	"In comparison to other optimal growth models… MERGE (Manne, Mendelsohn and Richels, 1995) links a simple top-down model to a bottom-up part that returns the cost of energy; in contrast, WITCH is a single model that represents the energy sector within the economy, and therefore chooses the energy technology investment paths coherently with the optimal growth structure. Also, WITCH features a non-cooperative game among the regions." (Bosetti, Carraro, Galeotti, et al., 2006, p. 16)
MERGE-ETL	MERGE	Version	"A MERGE model with endogenous technological progress" (Kypreos & Bahn, 2003, p. 249).
REMIND	MERGE WITCH	Structural	"With MERGE and WITCH, REMIND-R shares the same intertemporal structure, but is distinguished from both by a higher degree of technological resolution in the energy sector…Whereas WITCH is more elaborated in modeling R&D investments and knowledge spillovers, REMIND-R is more advanced in addressing trade issues" (Leimbach et al., 2010, p. 157).
BET	MERGE & MARKAL-MACRO	Version	According to Yamamoto et al. (2014) (2014, p. 584), BET "is strongly influenced by MERGE (Manne et al. 1995; Richels and Blanford 2008) and MARKAL-MACRO (Loulou et al. 2004)" (2014, p. 584) and "the BET model can be summed up as "a MERGE with advanced, electric end-use technologies" or "a global MARKAL-MACRO with limited technologies"" (Yamamoto et al., 2014, p. 585). In the acknowledgements, Yamamoto et al. write "we greatly appreciate the kindness of the MERGE group to make a version of the code available online, which helped us develop the BET model" (Yamamoto et al., 2014, p. 595).
MESSAGE-MACRO	MERGE + MESSAGE	Combination	"MACRO, as it is used in the link with MESSAGE, has its roots in a long series of models by Manne and others. The latest model in this series is MERGE–3" (Messner & Schrattenholzer, 2000, p. 271).

	MARKAL-MACRO	Similar Component	"…the MARKAL–MACRO model…has much in common with our model. The main difference between the two approaches is that MARKAL–MACRO is a fully integrated single model, whereas MESSAGE–MACRO is solved by running each part separately and iterating their inputs until consistency between the macroeconomic part and the energy part is reached." Messner & Schrattenholzer (2000, p. 270).
MARKAL-MACRO	MACRO + MARKAL	Combination	"MARKAL-MACRO is an experiment in model linkage" (Manne & Wene, 1992, p. 1)
AIM-Enduse	ERB	Component	"The energy sector top-down module was developed based on the revised Edmonds-Reilly-Barns (ERB) model (Edmonds et al. 1983; Edmonds et al. 1995)" (Kainuma et al., 2003, p. 20) †.
Phoenix	ERB	Version	"Phoenix is a re-design of the Second Generation Model (SGM) produced by The Joint Global Change Research Institute" Wing et al. (2011). "the Second Generation Model (SGM)…is a process-level version of the ERB formulated as a general-equilibrium energy-economy model"(Brenkert et al., 2003, p. 13) º.
TIMER	POLES	Similar	"…a model which is in various aspects similar to the TIMER-model is the POLES-model" Vries et al. (2001, p. 11).

† This is the only statement in the table that is not taken from the earliest model documentation. Early references for AIM and AIM-Enduse (such as Matsuoka et al. (1995) and Kainuma et al. (1995)) are either not available in English, or do not mention other models. The source here is a more recent book (published in English) about the AIM model, which also includes a detailed history of the model.

º Brenkert et al. (2003) provides model documentation for MiniCAM, not Phoenix. Two of the three authors, however, also wrote the model documentation for the SGM (Brenkert et al., 2004) – which is what became Phoenix – one year later.

In short, IAMs are linked if they are constructed by combining other models, or if they share the same structure as or a component with other models. IAMs are also linked if the model documentation describes them as versions of another model or indicate that they are "similar to" another model.

Some types of links are clearer in meaning than others. It is generally clearer what is meant by a 'Combination' link than by a 'Version' link. And it is generally clearer what is meant by a 'Version' link than by 'Similarity'. At the same time, however, it is important to understand that the typology is based on the wording of the model documentation, which reflects subjective judgments. Care should therefore be taken in drawing conclusions from the types of links. Whether a new model is described as a version of an older model and the link thus ends up being classified as a 'Version' link, or whether it is described as sharing a component and the link ends up being classified as a 'Component' link, or whether it is described as sharing the same structure and the link ends up being classified as 'Structural' link is to a large extent up to the modelers writing the model documentation. Thus, similar model relationships might end up being classified as different types of links.

In addition to this, the identification of links depends on what can be discerned from the model documentation. The amount and quality of documentation for different IAMs varies substantially. Some modelers provide comprehensive descriptions, comparisons, and histories of their model, others do not. In general, more influential IAMs are better documented. This means that model links are more likely to be identified for more influential IAMs. This introduces a potential bias in the findings. The documentation (available in English) for the two IAMs that don't have any links to other IAMs in figure 2.3 (IMACLIM and DNE21+), for example, was relatively limited. Appendix A provides more information about the model links, including the availability of model documentation.

Lastly, it is important to note that we cannot conclude from the types of links the strength of links. For our purpose, the strength of a link must be related to the degree of model dependence. The degree of model dependence will be higher if two models share more assumptions and model choices. In the end, however, dependence matters because it leads to correlated – and thus not independent – model results. The more highly correlated outputs are, the stronger a link can be said to be. But every IAM has multiple outputs, and different outputs are most likely affected differently by the same link. The use of carbon dioxide removal (CDR) in scenarios that limit global warming to 2°C, for example, might be very similar in two IAMs due to similarities in cost assumptions and the solution algorithm that determine technology choice. The same two IAMs might, however, provide very different estimates for the employment effects of reaching the same target, as this will be determined by other factors. Thus, the strength of a link is not a one-dimensional concept. And to determine the impacts of a link even for one particular model output, one would have to know first all the factors that determine the output in question and second how these factors are affected by the link. Due to the variety of relationships that are captured by the links and the subjectivity involved in describing these, we cannot infer from the types of links the impacts of links on outputs. Determining the impacts of links on outputs would require a detailed and comprehensive comparison of IAMs and their outputs.

2.6.2 From links to dependencies

We do not, however, need to know the strength of links or the impacts of links on outputs in order to argue that the links depicted in figure 2.3 imply *likely model dependencies*.

One of the reasons why it is difficult to conclude from the types of links the strength of links is that different types of links can represent similar relationships. A new version of an older model, for example, would normally inherit the structure as well as components from the older model. The link between these models might thus be classified as a 'Version', 'Component', or 'Structural' link, depending on how modelers choose to describe the relationship in the model documentation. The fact that different types of links can represent similar relationships, however, is an important point in itself. A 'Version' link indicates that either the core of the model is the same (such as is the case for MERGE and MERGE-ETL), or that the newer model inherits important building blocks (such as is the case for Phoenix) or code (such as is the case for BET) from the older model. For 'Component' links, a part of the older model is used in the newer model. For 'Structural' links, the model structure is the same. In other words, the reason why it is difficult to distinguish between 'Version', 'Component', and 'Structural' links is that they all imply that either model components and/or structure is shared. The same is true for 'Combination' links.

Thus, we find that all the link types except 'Similar' links, imply shared model components, shared model structure, or both. In all these cases, certain model choices and assumptions will be shared. Thus, all the link types – except for one – indicate some degree of model dependency. This might also be the case for 'Similar' links, but the statements on which these links are based are too vague to warrant conclusions. It is, however, reasonable to think that when modelers state that two models are "in various aspects similar" this also includes some model choices and assumptions. Still, the 'Similar' link type is only used on its own in one instance (between POLES and TIMER).

Thus, we can say, without knowing the strength or the specific impacts on model outputs, that (almost) all the links in the model family tree imply likely model dependencies. Despite the vagueness of many of the statements on which the links are based, the dependence conclusion is further supported by the fact that model documentation in most cases is highly selective about what other IAMs are referenced. In no cases do the model documentation for a given IAM compare the IAM against "all" other IAMs. In other words, even though the links in some cases might appear uncertain or even tenuous, the fact that another model is mentioned in the model documentation for an IAM in the first place is itself an indication of significance.

One might, however, argue that being mentioned is an indication of the significance of the model being mentioned, not the link. By discussing another IAM, this IAM is given space, which is an indication that this IAM is seen as significant or somehow relevant to the authors. In particular, one would expect better-known and more highly regarded IAMs to be seen as more significant to all modelling groups and thus mentioned more often. Still, there are no examples of IAMs that are mentioned in the model documentation for all the 14 IAMs. Even though MERGE is a well-known and long-standing model – something which might explain part of the reason why model documentation so often refers to it – it is important to realise that far from all IAMs in AR5 mention MERGE. Only a distinct group of IAMs do.

Therefore, one might alternatively take the branches in the model family tree to indicate *social scientific networks*. IAMs within the same network are considered more significant and therefore relevant to other IAMs within the same network, which make them more likely to mention each other. If we take links to indicate social scientific networks, the branches in the model family tree can be seen to represent distinct networks.

What does that imply for the dependence conclusion? Everything else being the same, two models in the same network are more likely to learn from each other and therefore make similar model choices and assumptions than two models belonging to different networks. In other words, even under this interpretation, the links can be seen to indicate closer relationships, and thus likely model dependencies.

Overall, even though the types of links cannot be used to infer the strength of the links or the impacts on different outputs, they can be used to suggest that linked IAMs are not independent, and thus that some of their outputs are likely to be correlated.

2.7 The evolution of IAMs

In addition to constructing the model family tree, the model documentation was used to obtain a richer understanding of how the 14 IAMs have evolved as a group.

2.7.1 Increasing detail and scope

All IAMs that have been around for some time have gone through multiple versions. For each new version, changes are made. The model documentation for IMAGE, which is one of the most well documented IAMs, provides a detailed illustration of this process. IMAGE has been developed progressively since the 1980s with the inclusion of additional systems (e.g. energy and agriculture), regional detail, improved representations of systems (e.g. climate and land use), and other additional details (Stehfest et al., 2014). Following the initial version (IMAGE 1.0), “IMAGE 2.0 was the first published global integrated model having geographic resolution” (Bouwman et al., 2006, p. 9). In the late 1990s, “further refinements and extensions were implemented in IMAGE 2.1…to enhance the model’s performance and broaden its applicability” (Bouwman et al., 2006, p. 10). Only one year later, “the board recommended making Global Change the target area, extending it beyond climate change, and building on integration of socio-economic and natural systems” (Bouwman et al., 2006, p. 10).

IMAGE 2.6, again, "marks an important milestone on the development path towards a next generation model, referred to as IMAGE 3, aimed at capturing – to a larger extent – the different aspects and domains of sustainability, with emphasis on the ecological domain but also related to the economic and social domains" (Bouwman et al., 2006, p. 17). In summary, IMAGE has been continuously expanded over time to increase the detail and scope of analysis.

Although IMAGE is the best documented of the 14 IAMs, we see similar trends when we examine the model documentation also for the other IAMs. For GCAM, the increase in scope and model capabilities over time has been a response to a growing set of questions: "throughout its lifetime, GCAM has evolved in response to the need to address an expanding set of science and assessment questions. The original question that the model was developed to address was the magnitude of mid-21st-century global emissions of fossil fuel CO_2. Over time GCAM has expanded its scope to include a wider set of energy producing, transforming, and using technologies, emissions of non-CO_2 greenhouse gases, agriculture and land use, water supplies and demands, and physical Earth systems" (GCAM v5.1 Documentation, 2019).

Similarly, MERGE and WITCH (even though the latter has only been around since 2006) also show signs of increasing scope and detail over time. According to Blanford et al. (2014, p. 528) "like virtually all models being actively used in the climate debate, MERGE is continually being adapted to assess the implications of new policy proposals. Among the most noteworthy enhancements to the current version is the inclusion of BECS, **b**io-energy with **c**arbon dioxide **c**apture and geologic **s**torage". POLES has also been "extended on several occasions to capture the most recent market and policy developments" (Despres, 2018).

Thus, the model documentation for several of the 14 IAMs suggests that policy is an important driver in IAM development. This appears to lead to a continuous increase in detail and scope. A similar observation was made in 2003, when Kainuma et al. (2003) (working with the AIM model) noted that the development of IAMs tended towards an inclusion of more phenomena (thus widening the scope) and more detail, and towards applications to regional and local scales, which required an increase in resolution. These developments were, according to Kainuma et al., a consequence of the increasingly central role that IAMs play at the interface of science and policy and the associated demands that are put on them. The expansion of scope might also be explained by the fact that it increases the audience and users of IAMs.

The tendency towards increasing scope and detail is not only observed in the evolution of individual IAMs over time. A similar trend is also found if one looks at how new IAMs distinguish themselves from existing IAMs. MARKAL-MACRO, for example, was constructed specifically to increase the scope and detail of analysis, again motivated by policy demands. As Manne and Wene wrote when they presented MARKAL-MACRO for the first time, “an efficient modelling tool must have the scope and detail to match the width and depth of the policy problem being analyzed” (Manne & Wene, 1992, p. 1). When it comes to the MESSAGE/MERGE branch, we find that increasing detail and resolution is frequently emphasised as the distinguishing feature for several of the new IAMs. REMIND, for example, is distinguished from MERGE by a higher degree of technological resolution in the energy sector and a higher resolution in the representation of trade and BET is seen as “a MERGE with advanced, electric end-use technologies” (Yamamoto et al., 2014, p. 585). The documentation for WITCH also emphasises “richer technological detail” (Bosetti, Carraro, Galeotti, et al., 2006, p. 16) relative to other IAMs (see Appendix A for more details). The hybrid nature of newer IAMs, i.e. the way in which they combine previously distinct types of IAMs[18] into a single framework, is sometimes also emphasised. REMIND, for example, was designed as a hybrid model from the start. For practical purposes, hybridisation makes IAMs more “complete” when it comes to their ability to analyse different questions.

2.7.2 A “normal” evolution

If the development of IAMs is anything like the development of climate models, there are reasons to believe that new IAMs are designed partly by combining and modifying existing parts and ideas. This is most obvious when two existing models are combined to construct a new one (thus forming a ‘Combination’ link). As Wene (1996) writes in his description of the linking between ETA-MACRO (later MERGE) and MESSAGE, “for linking, it is possible to use peer-reviewed models, which avoids repeating earlier work and provides needed initial quality assurance to the efforts” (p. 810). In other cases, and as shown by Knutti et al. (2013) for climate models, dependencies often arise due to the use of similar approaches and simplifications.

In many ways, the model family tree (figure 2.3) is suggestive of a research field that has evolved from a formative stage (prior to the 1990s), through an expansive stage (the 1990s), and to a mature stage (post 2000s). The observed trend among the 14 IAMs towards an increase in detail and scope over time also fits with the Kuhnian concept of ‘normal science’ (Kuhn, 1962). ‘Normal science’ refers to a stage that is characterised by incremental progress and a continuous accumulation of detail within established frameworks. In this stage, underlying assumptions are not questioned (Kuhn, 1962).

[18] Combining what has been referred to in the literature as “top-down” (primarily macroeconomic) models with “bottom-up” (typically engineering based) models.

A close reading of the model documentation for the IAMs constructed after 2000 shows that the underlying assumptions made in comparable older IAMs are not challenged or questioned. Structural changes or the development of new approaches are also not emphasised. While the documentation for two of the IAMs in the MESSAGE/MERGE branch, MERGE-ETL and WITCH, emphasise the endogenous representation of technological change[19] – which in certain respects (in particular for the cost of mitigation) represents a significant departure from models with exogenous technical change – this is still done by modifying existing approaches (based on optimal growth theory). If anything, comparisons of new IAMs to old IAMs that do not challenge any of the model choices or assumptions provide an implicit indication that new IAMs are based on old approaches.

This, and the fact that new IAMs are distinguished from older IAMs primarily by pointing to an increase in detail, suggest a trend in the evolution of IAMs towards an expansion of what is, rather than towards a diversity of approaches. Like most mature research fields, the development of integrated assessment modelling, at least post 2000, appears to be path dependent and incremental. Bosetti et al., for example, "wrote the first equations of WITCH in 2005 and since then the model has grown in complexity and richness. However, the core of the model has not changed" (2014, p. viii). While such a trend might represent a natural progression of a research field, which is not a problem in itself, it may limit the structural diversity of IAM ensembles. The development of IAMs has been driven to a large extent by policy demands, which have led IAMs to attempt to answer an expanding set of questions at an increasing level of resolution. This has incentivised an increase in detail and scope, but not necessarily a variety of assumptions and modelling choices. Unless structural diversity is encouraged specifically, policy relevance is more likely to define the development of IAMs, also in the future[20]. This might pose problems if IAM ensembles are used in the belief that they capture the relevant structural diversity.

2.8 Conclusion

This chapter has argued, based partly on studies published in the climate modelling literature, that IAM dependencies is an important but neglected topic in AR5. If IAMs are not independent, we cannot know whether agreement in results is a sign of robust insights or a consequence of shared model choices and assumptions. In order to identify likely model dependencies in the AR5 IAM ensemble, the chapter has

[19] ETL in MERGE-ETL stands for endogenous technical learning and the development of WITCH was motivated by a need to "capture the dynamics of technical change and the relationships between technical change and the main economic and policy variables" (Bosetti et al. in 2006 p. 14). For WITCH, "the endogenous representation of R&D diffusion and innovation processes constitute a distinguishing feature" (The WITCH team, 2017, p. 5).

[20] The discussion here is limited by what could be discerned from the model documentation. There are also other incentives, such as academic incentives, that are likely to affect the evolution of IAMs.

developed a method that deliberately avoids comprehensive investigations of individual IAM assumptions. The method, which is based on analysing available model documentation, culminates in a model family tree that shows when the most influential IAMs in AR5 were designed and how they are linked to each other. The model family tree shows that the 14 most influential IAMs in AR5 form three distinct branches of IAMs, the largest of which is the MESSAGE/MERGE branch. Together, the IAMs in the latter branch (MERGE, MESSAGE-MACRO, MERGE-ETL, REMIND, WITCH, and BET) are responsible for about half of the scenarios in the AR5 ensemble. By analysing the types of links, it is shown, all the links in the model family tree (except for possibly one) imply that assumptions and model choices are shared to some extent. Thus, IAMs that are linked are not independent. In addition to this, the model documentation is used to provide an understanding of what has driven the development of IAMs over time. It is found that the development of IAMs has been driven, to a large extent, by policy demands that have incentivised increasing detail and scope. If we consider not only what is stated, but also what is not stated in the model documentation, we find that new IAMs rarely challenge the model choices and assumptions that are made in related IAMs.

The extent of the IAM dependencies shown to be present in the AR5 IAM ensemble presents two main issues. First, the extent of the dependencies challenges the robustness of AR5 insights. In particular, the findings suggest, agreement among IAMs in the MERGE/MESSAGE branch could be a result of shared assumptions and modelling choices rather than a sign of robustness. Second, the extent of the dependencies – which imply shared model choices and assumptions – limits the structural diversity of the AR5 IAM ensemble. This implies that the AR5 IAM ensemble might not provide a good representation of the structural uncertainty that is associated with IAM outputs. The severity of these two issues for AR5 results, however, can only be determined by conducting more detailed analyses of the effects of the model dependencies on specific IAM outputs.

The method developed in this chapter circumvents hugely time-consuming model comparisons in order to enable identification of model dependencies in an entire ensemble of IAMs. While the results indicate that model dependencies are present among several highly influential IAMs in AR5, more work is needed in order to determine the strength and implications of these dependencies. For this purpose, a number of approaches might be taken.

Jun et al. (2008) showed that institutional links have a significant impact on climate model dependencies. One thing that is not mentioned in Table 2.3 (but which can be discerned from Table 2.2) is that the model that later became Phoenix, SGM, was developed by some of the same researchers who developed the model that later became GCAM, MiniCAM, at the same institution (the Pacific Northwest National Laboratory (PNNL)). This adds further support to the claim that Phoenix and

GCAM are not independent.[21] In order to gain more information about the institutional and personal links that shape IAM development and research, bibliometric and social network analysis could be used. Examples of relevant studies include Chappin et al. (2014), who use bibliometric analysis to study the structure of authorship networks in the field of socio-technical transitions research, and Corbera et al. (2016), who employ social network analysis to study patterns of authorship in the AR5 WGIII report. Applying similar techniques to IAM publications could be used to provide additional information regarding the links between IAMs. In particular, metrics from social network analysis that capture centrality and clustering might be used to construct measures of IAM relatedness. Such metrics would still, however, be limited in that they can only serve as proxies for dependencies in model choices and assumptions. It is still possible that even highly connected IAMs (in terms of people and institutions) differ when it comes to important model choices and assumptions.

In order to get a better understanding of the strength and implications of the model links identified in this chapter, interviews with modelers could therefore be used. One might ask questions such as: "In what ways is model X similar to model Y and in what ways is it different?" and "What are the implications for results of the link between model X and Y?". One could also use interviews to ask more questions about the history of the models (who developed them, on what basis, and for what purpose). The information that can be obtained from interviews is, however, limited by the knowledge and viewpoints possessed by present-day modelers. The model documentation instead gives us the viewpoints of the modelers who designed the models at the time of construction. For some of the most long-standing IAMs, present-day modelers may be less aware of the original roots and connections to other models. Interviews might thus be more informative in the case of newer IAMs. The main limitation to interviews in this case is the possibility that model developers may want to highlight novelty and end up underplaying similarities and model dependencies. This, however, may also be true for the model documentation.

A richer account of the evolution and development of IAMs could also be obtained by conducting a more comprehensive historical analysis. Such an analysis could even investigate the process by which models and scenarios have been developed and ultimately included in IPCC reports. The material used for this purpose should be expanded beyond model documentation to include other sources that may provide information regarding IAM development, key institutions (such as for example IIASA and the Stanford Energy Modelling Forum), influential modelers, and the processes by which IAMs are included in IPCC reports. Among other things, such an analysis might help explain why all the IAMs

[21] The reason why this was not discussed in the main body of this chapter is that institutional and personal links were not investigated in a systematic manner.

that have been developed post 2000 belong to the MERGE/MESSAGE branch. It might also help explain why an IAM such as E3MG, which the modelers claim offers a fundamentally different approach compared to other IAMs (Barker et al., 2012; Barker & Şerban Scrieciu, 2010), was included in AR4, but not in AR5.

In order to support, or challenge, the significance of the IAM dependencies that have been identified in this chapter, a mostly data centred approach could be used to determine whether the IAMs that are linked in the model family tree produce correlated outputs. The WGIII AR5 Scenario Database is available online (IAMC, 2014), and so are other databases that contain model results for many of the IAMs analysed in this chapter. Data from these databases could be used to investigate whether model dependencies imply similarities in results.

Awaiting further research into the strength and the implications of the likely IAM dependencies that have been identified in this chapter, the IPCC could still take a number of steps to improve the interpretation of ensemble results and ameliorate the potential consequences of IAM dependencies. The presence of IAM dependencies should not come as a surprise, given how modelling work is known to be conducted, and given how climate models have been shown to relate.

The first thing the IPCC can do is acknowledge and discuss IAM dependencies in the presentation of ensemble results. One of the core aims of the WGIII reports is assessing the uncertainty of the knowledge base related to climate mitigation. By not considering IAM dependencies, it might be argued that AR5 fails to meet this aim regarding transformation pathways, which represent key contributions to the AR5 report. The acknowledgement of IAM dependencies should therefore be done in the outer layers of the IPCC reports, such as the SPMs and the Synthesis Reports.

The second thing the IPCC can do is improve the presentation of ensemble results by addressing the potential impacts on results from IAM dependencies directly. The uneven number of scenarios generated per IAM in AR5 (shown in Figure 2.2) – which is mentioned in passing in Chapter 6 of the AR5 WGIII report – means that scenarios produced by some IAMs (especially the Big Four) have a significant influence on ensemble results. The number of scenarios produced by dependent IAMs, however, is still much more numerous. In order to ameliorate the unevenness of scenarios generated by each IAM, IPCC reports could simply weigh each IAM, rather than each scenario, equally[22]. In order to ameliorate the dominance of certain branches of IAMs – which is the focus of this chapter – the IPCC

[22] Under this scheme, each REMIND scenario would have a weight of 1/(158*30) whereas the single Ecofys scenario would have a weight of 1/30.

reports could weigh each scenario according to the branch it belongs to such that each branch of IAMs, rather than each scenario, is given equal influence on ensemble results. Another benefit of so doing is that it would illuminate the extent to which model dependencies actually matter for results. Based on different ways of grouping dependent IAMs, different results might be obtained. If the groupings have a big impact on ensemble results, a more detailed investigation of dependencies would be warranted. If the groupings don't make a big difference, less attention would need to be paid to the dependencies. In the end, this brings us back to the data analysis suggested above for investigating correlations between outputs from IAMs.

The third thing the IPCC can do is take steps that might lead to an increase in the diversity of future IAM ensembles. As Masson & Knutti argue, "the goal for an ensemble should be to maximize diversity … and minimize dependency" (2011, p. 3). Among other things, the goal of maximising diversity could be built into the process by which IAMs are selected for inclusion in IPCC reports. If the problem instead is a scarcity of diverse IAMs in the literature, the IPCC could also use its authority to call for the development of a wider variety of IAMs. The IPCC could also call for more research on IAM dependencies. By so doing, more attention would be paid to those IAMs that bring independent insights, which again might incentivise the development of such IAMs.

Overall, this chapter has taken a first step towards assessing some of the important challenges associated with the use of IAM ensembles in IPCC reports. It has done so by drawing on relevant studies in the climate modelling literature and by developing a method for identifying likely model dependencies in the AR5 IAM ensemble. One of the key challenges associated with studying IAM ensembles as distinct entities is that the number and complexity of IAMs in ensembles preclude a deep comparison of individual IAM features. This means that the analysis in some ways might appear shallow. However, unless we study IAM ensembles *as ensembles*, we might fail to recognize some of the overarching conditions that are necessary for our ability to obtain robust insights into climate mitigation. Although the method developed in this chapter does not obviate the need for more detailed analyses of the implications of IAM dependencies, it has shown how such dependencies are likely to appear among IAMs and it has identified a number of dependencies that can serve as the starting point for more detailed investigations.

Chapter 4 will look in more detail at how certain kinds of IAMs in AR5, many of which are found in the MERGE/MESSAGE branch, might have led to a bias in the reported mitigation costs. Before then, the next chapter presents the dimensions that are used to classify IAMs in the (appendix to the) WGIII report and the frameworks that the IAMs in AR5, according to the model documentation, are based on. This is used to show that the branches in the model family tree, although they align with underlying model frameworks, also provide information about model dependencies and independencies that are not captured by looking at IAM frameworks alone.

3 Model Frameworks

Chapter 2 shows that several of the more recently developed IAMs in the AR5 ensemble share either model structure or model components (or both) with older IAMs. Although AR5 does not discuss IAM dependencies, it does contain a general discussion of key IAM differences. This chapter employs the information provided in the AR5 WGIII report appendix regarding these differences in order to categorise the 30 IAMs in the AR5 ensemble. The resulting categories, which reflect the degree of foresight and economic coverage in IAMs, are then compared with the model frameworks that are used to classify the 14 most influential IAMs in the respective model documentation. The chapter shows that there is an almost perfect match between the degree of foresight and economic coverage and the underlying frameworks. It also shows that there is a strong overlap between the categories and model frameworks, and the model family tree branches identified in Chapter 2. In particular, all the IAMs in the MESSAGE/MERGE branch are general equilibrium – perfect foresight models based on a Ramsey-type optimal growth framework. Still, however, the method developed in Chapter 2 captures sources of model dependencies and independencies that the information contained within the WGIII appendix does not capture.

While the model categories and frameworks presented in this chapter will be familiar to most modelers and many researchers in this area, the uneven distribution of IAMs in AR5 with respect to these frameworks might not be. This is a point that will be brought up also in Chapter 4, which highlights additional reasons for why diversity in IAM ensembles is important.

Section 3.1 presents the six “key differences in model structure” that are discussed in the WGIII report and uses two of these (the degree of foresight and economic coverage) to group the AR5 IAMs. The results are also compared to the model family tree constructed in Chapter 2. Section 3.2 adds to this the model frameworks that, according to the model documentation, underpins the 14 most influential IAMs in AR5. Section 3.3 further discusses the relationships between the model family tree constructed in Chapter 2 and the model differences and frameworks analysed in this chapter and provides concluding remarks.

3.1 Key structural differences

All IAMs in AR5 fall under the generic category of ‘large-scale integrated models’. While AR5 highlights the role of integration and interdisciplinarity in IAMs, it provides little information about what exactly is integrated and how. The most detailed explanation of what IAMs do is found in Chapter 6 of the AR5 WGIII report, which states that

> “[IAMs] use economics as the basis for decision making. This may be implemented in a variety of ways, but it fundamentally implies that the models tend toward the goal of minimizing the aggregate economic costs of achieving mitigation outcomes, unless they are specifically constrained to behave otherwise. In this sense, the scenarios tend towards normative, economics-focused descriptions of the future. The models typically assume fully functioning markets and competitive market behavior, meaning that factors such as non-market transactions, information asymmetries, and market power influencing decisions are not effectively represented.” (IPCC, 2014a, p. 422).

This description, however, still leaves a lot unanswered. Economics can be used in many different ways to describe decision making and it is not clear what “economics-focused descriptions of the future” mean. Even if the economics framework was clear, the above quote only describes *tendencies* and *typical* assumptions. It appears that IAMs can always be constrained to behave differently. Overall, the above description provides only a vague understanding of what IAMs do and what they can and cannot tell us. This might not be surprising given the complexity and size of IAMs, which makes it very difficult to make general claims.

Chapter 6 does, however, highlight six “key differences in model structure” between IAMs (IPCC, 2014a, p. 422):

1. **“Economic coverage and interactions”.** This captures the way in which models “differ in terms of the degree of detail with which they represent the economic system and the degree of interaction they represent across economic sectors” (IPCC, 2014a, p. 422). Two options are contrasted: “*Full-economy models* (e. g., general equilibrium models) represent interactions across all sectors of the economy, allowing them to explore and understand ripple effects from, for example, the imposition of a mitigation policy, including impacts on overall economic growth. *Partial-economy models*, on the other hand, take economic activity as an input that is unresponsive to policies or other changes such as those associated with improvements in technology. These models tend to focus more on detailed representations of key systems such as the energy system” (IPCC, 2014a, p. 422). Note that, in the appendix to the WGIII report, the more common term *partial-equilibrium* is used instead of *partial-economy*.
2. **“Foresight”.** Again, two options are contrasted: “*Perfect-foresight models* (e. g., intertemporal optimization models) optimize over time, so that all future decisions are taken into account in today’s decisions. In contrast, *recursive-dynamic models* make decisions at each point in time based only on the information in that time period” (IPCC, 2014a, p. 422). The latter models are also referred to as *myopic* models in the WGIII report.

3. **“Representation of trade”.** This reflects the fact that “models differ in terms of how easy it is for goods to flow across regions” (IPCC, 2014a, p. 422).
4. **“Model flexibility”.** According to Chapter 6 of the WGIII report, “the flexibility of models describes the degree to which they can change course. Model flexibility is not a single, explicit choice for model structure. Instead, it is the result of a range of choices that influence, for example, how easily capital can be reallocated across sectors including the allowance for premature retirement of capital stock, how easily the economy is able to substitute across energy technologies, whether fossil fuel and renewable resource constraints exist, and how easily the economy can extract resources” (IPCC, 2014a, p. 423).
5. **“Sectoral, regional, technology, and GHG detail”.** According to Chapter 6, “Models differ dramatically in terms of the detail at which they represent key sectors and systems…Key choices include the number of regions, the degree of technological detail in each sector, which GHGs are represented and how, whether land use is explicitly represented, and the sophistication of the model of earth system process such as the carbon cycle” (IPCC, 2014a, p. 423).
6. **“Representation of technological change”.** Here, two options are again contrasted: “On one end of the spectrum, models with *exogenous technological change* take technology as an input that evolves independently of policy measures or investment decisions. These models provide no insight on how policies may induce advancements in technology. On the other end of the spectrum, models with *endogenous technological change* (also known as *induced technological change*) allow for some portion of technological change to be influenced by deployment rates or investments in research and development (R&D)” (IPCC, 2014a, p. 423).

These key differences offer useful insights into the workings of IAMs and some of the trade-offs that are associated with integrated assessment modelling. The WGIII report does not, however, relate any of these differences to the IAMs in the AR5 ensemble, except for in the appendix. More specifically, the WGIII report appendix includes a table (Table A.II.14) that lists certain features of the 30 IAMs that are included in the AR5 ensemble. Among other things, this table classifies each IAM according to three out of the six key differences described above: economic coverage, degree of foresight, and level of detail (for regions and GHGs). No information is provided in AR5 for the IAMs regarding the remaining three key differences (representation of trade, model flexibility, and representation of technological change). Given our interest here in the frameworks that underpin IAMs rather than the levels of detail (which Chapter 2 showed tends to increase over time), this leaves us with two dimensions according to which AR5 IAMs can be classified using the information in the appendix: economic coverage and degree of foresight. The appendix provides three options for economic coverage – *general equilibrium*, *partial equilibrium*, and *econometric* – and two options for the degree of foresight – *myopic* and *foresight*. (While AR5 provides a brief description of the differences between general and partial equilibrium models, econometric models are not discussed.) This allows us to construct a 2x3 matrix in which to place the 30 IAMs, which gives rise to six categories of IAMs. This is shown in Table 3.1.

Table 3.1 Classification of AR5 IAMs according to the two AR5 dimensions, 'Degree of Foresight' and 'Economic coverage'. The 14 most influential IAMs (analysed in Chapter 2) are shown in black and the remaining IAMs are shown in grey. The fractions of scenarios generated by all the IAMs in each of the six categories as a proportion of the entire AR5 ensemble are shown in parenthesis.

	General equilibrium	Partial equilibrium	Econometric
Foresight	REMIND, MESSAGE-MACRO, WITCH, MERGE-ETL, MERGE, BET EC-IAM, GRAPE, MARIA, iPETS **(50%)**	DNE21+, TIAM-World, TIAM-ECN, TIMES-VTT **(9%)**	IEEJ **(0.2%)**
Myopic	IMACLIM, Phoenix ENV-Linkages, FARM, GEM-E3-ICCS, WorldScan2, SGM, IGSM, GTEM, KEI-Linkages **(13%)**	GCAM, IMAGE, POLES, AIM-Enduse Ecofys Energy Model **(29%)**	

Table 3.1 shows two things. First, it shows that some IAM categories are responsible for significantly more scenarios in the AR5 scenario ensemble than others. The most influential IAM category is the *General equilibrium - Perfect foresight* category, which is responsible for 50% of the scenarios in AR5. The second largest category is the *Partial equilibrium - Myopic* category, which is responsible for 29% of the scenarios. Scenarios from the *Partial equilibrium – Perfect foresight* and the *General equilibrium – Myopic* categories account for only 9% and 13% each. The *Econometric* category includes only one IAM, which is responsible for only 2 scenarios. Thus, the influence of this category on the AR5 ensemble is insignificant. Second, it shows that the IAM categories in Table 3.1 have a strong overlap with the branches in the model family tree presented in Chapter 2. The six IAMs (of the 14 most influential, shown in black) in the upper left corner of the table show a perfect correspondence with the six models that make up the 'MESSAGE/MERGE branch' in the model family tree. And almost all the IAMs that constitute the two remaining branches of the model family tree (POLES and IMAGE on the one hand, and AIM-Enduse and GCAM on the other hand) fall into the *Partial equilibrium - Myopic* category. There is only one instance where two models belong to the same branch in the model family tree but fall into separate categories in Table 3.1, namely GCAM and Phoenix.

3.2 Underlying model frameworks

The previous section places the 30 IAMs according to some of the key differences in model structure listed in AR5. AR5 does not, however, relate these structural differences to the frameworks that are typically used in the literature to describe IAMs. In order to find this information, we again look to the model documentation.

The frameworks that, according to the model documentation, underpin the 14 most influential IAMs in AR5 are shown in Table 3.2 (in orange) together with the categories identified in the previous section. In short, models based on a *Ramsey-type optimal growth* framework compute an optimal pathway by choosing the savings and investments that optimise the intertemporal utility of a representative agent with perfect foresight (Rezai et al., 2013). Similarly to these models, *computable general equilibrium (CGE)* models also compute optimal pathways arising from the optimising behaviour of agents (Peace & Weyant, 2008), but whereas optimal growth models treat the entire economy as one sector, CGE models include the interactions between different economic sectors. Chapter 4 will discuss the implications of using optimal growth models and CGE models to estimate the cost of mitigation. Lastly, *energy system optimisation models (ESOMs)*, instead of optimising the utility of a representative agent, optimise the total energy system cost (or, equivalently, the total surplus in energy markets). ESOMs use linear programming, which is a special case of mathematical programming, to identify the optimal solutions. Some of the consequences of using this technique to compute transformation pathways will be discussed in Chapter 5. More details regarding the descriptions in the model documentation of what frameworks are used in each of the 14 IAMs are provided in Appendix B.

The main takeaway from Table 3.2 is that the categories that arise based on the two dimensions, 'Degree of Foresights' and 'Economic coverage', fit almost perfectly onto the model frameworks used to describe the 14 IAMs in the model documentation. We already saw that the six models that make up the 'MESSAGE/MERGE branch' in the model family tree (MERGE, REMIND, WITCH, MERGE-ETL, BET, and MESSAGE-MACRO) all fall into the same category in Table 3.1 (*General Equilibrium - Perfect Foresight)*. Table 3.2 shows that all of these models are based on a *Ramsey-type optimal growth* framework. (MESSAGE-MACRO simply combines two frameworks, corresponding to MESSAGE and MACRO respectively). The two models in the *General Equilibrium* - *Myopic* category (IMACLIM and Phoenix) are both CGE models and the two models in the *Partial Equilibrium* - *Perfect Foresight* category (DNE21+ and TIAM-World) are both ESOMs.

Table 3.2 Classification of the 14 most influential AR5 IAMs according to the two AR5 dimensions, 'Degree of Foresight' and 'Economic coverage' with additional information from the model documentation regarding the model frameworks. The number of scenarios generated in total by all the IAMs in each of the four categories are shown in parenthesis.

	General equilibrium	**Partial equilibrium**
Foresight	*Ramsey-type optimal growth* REMIND, WITCH, MERGE-ETL, MERGE, BET *Ramsey-type optimal growth + Energy system optimisation (ESOMs)* MESSAGE-MACRO **(545)**	*Energy system optimisation (ESOMs)* DNE21+, TIAM-World, **(84)**
Myopic	*Computable general equilibrium (CGE)* IMACLIM, Phoenix **(84)**	IMAGE, POLES, AIM-Enduse, GCAM **(338)**

The almost perfect correspondence between the frameworks described in the model documentation and the two AR5 dimensions used to categorise the IAMs in Table 3.1 might not be surprising. In fact, the key differences listed in AR5 are most likely based on these frameworks in the first place. It is worth noting, however, that it only takes two dimensions (and two options in each) to capture the main frameworks used to described IAMs. Clearly, these two dimensions are important in the field of integrated assessment modelling.

By taking a closer look at the actual frameworks, additional model similarities can also be found. In particular, all of the models in these three categories are based on optimisation algorithms of various sorts. Furthermore, the link between Ramsey-type optimal growth models and ESOMs appear to be relatively strong. Both MESSAGE-MACRO and MARKAL-MACRO – both of which are highly influential (in AR5 and beyond) – are based on linking an optimal growth model with and ESOM. Essentially, perfect foresight is intimately related to inter-temporal optimisation algorithms, which are used in both model types. This might make combining models more straightforward. The close relationships between these two model types is also confirmed by Nordhaus (2017), who writes that optimal growth IAMs and ESOMs share a long and mutually influential history dating back to at least the 1970s.

The frameworks used for the IAMs in the *Partial Equilibrium* - *Myopic* category are much harder to pin down. In a sense it is, perhaps, the most diverse category. This might explain part of the reason why many IAMs that fall into this category in Table 3.1 are not linked in the model family tree. IMAGE, which is described as closer to earth systems models than any other IAM, certainly appears quite unique in the IAM ensemble. What we *can* see, however, is that the models in this category tend to focus on *simulation* as opposed to *optimisation.* This generally means that they are based on algorithms that compute solutions at each point in time based on the state of the modelled system at the previous point in time together with assumptions about how the system (or agents within the system) behaves and responds to changes[23]. While very few IAMs claim to be able to forecast the effects of energy and climate policies, some of the simulation models, such as POLES (Enerdata, 2019), do. The distinction between optimisation and simulation will be discussed further in Chapter 5. Because these models focus mainly on the energy system (although GCAM also has a broader economic coverage), the term ‘energy simulation models’ is used to refer to them in this thesis.

3.3 Conclusion

This chapter has used the information regarding “key differences in model structure” presented in the WGIII report appendix to categorise the AR5 IAMs and compared the resulting categories with the frameworks that, according to the model documentation, underpin the 14 most influential IAMs. This, again, was compared with the model family tree branches that were identified in Chapter 2 of this thesis. The chapter has shown three things. First, it has shown that the information provided in the WGIII appendix is sufficient to delineate the main model frameworks that are used to describe the most influential IAMs in AR5: By categorising each of the 14 IAMs according to economic coverage and degree of foresight, Ramsey-type optimal growth models, Energy system optimisation models (ESOMs), and Computable General Equilibrium (CGE) models emerge as distinct groups (Table 3.2).

Secondly, this chapter has shown that the IAMs in the AR5 ensemble are unevenly distributed among the model categories that emerge. The *General equilibrium* - *Perfect foresight* IAMs are responsible for 50% of the scenarios in the AR5 ensemble. Of the 14 most influential IAMs that fall into this category, all are Ramsey-type optimal growth models. This gives reason to believe that the AR5 ensemble might be biased in the direction of model choices and assumptions implied by this framework[24]. The second largest category of IAMs, the *Partial equilibrium* - *Myopic* category, is

[23] This is in line with e.g. Edenhofer et al. (2006) who add that these models, mathematically speaking, solve initial or boundary value problems given a system of differential equations.

[24] Recall that the 14 most influential IAMs are responsible for 90% of the scenarios in the AR5 ensemble. Even if none of the remaining 16 IAMs in the AR5 ensemble (shown in grey in Table 3.1) are Ramsey-type optimal

responsible for 29% of the scenarios. This category, however, does not correspond to a distinct IAM framework according to the model documentation. Finally, the one IAM that is classified as an econometric model makes an insignificant contribution to the AR5 ensemble. AR5 does not recognise or discuss the uneven contribution of IAMs belonging to different categories. Measures could have been taken by the IPCC to present ensemble results in a way that reflects this unevenness. For example, and as already suggested in the conclusion to Chapter 2, each scenario could have been weighted in such a way that each category of IAMs is given equal influence on ensemble results (instead of each scenario).

Third, this chapter has shown that there is a relatively strong correspondence between the IAM categories that emerge based on the "key structural differences" (Table 3.1), the underlying IAM frameworks (Table 3.2), and the branches in the model family tree (Figure 2.3). With one exception, all the IAMs that belong to the same branch in the model family tree also fall into the same category when it comes to economic coverage and degree of foresight. In particular, of the 14 most influential IAMs, all the IAMs that fall into the *General equilibrium - Perfect foresight* category are the Ramsey-type optimal growth models that belong to the MESSAGE/MERGE branch. The overlap between the model family tree and the key structural differences also indicates that most of the model links in Chapter 2 either directly capture or are related to similarities in model structure.

For the 'Structure' links, this is not be surprising. This link type is used when the model documentation indicates similarity in structure and when the stated differences refer only to aspects such as model resolution or the way in which sub-model are combined. More specifically, Table 2.3 shows, this link type is used only for IAMs that are based on optimal growth theory. For the 'Combination' links, entire models are combined, and hence the structure is too. For the 'Component' links, structural similarity might either stem from the component that is being shared, or structural similarity might be a prerequisite for sharing components in the first place. When it comes to the 'Similarity' links, which are vague, it also makes sense that model structure, given its central role, is an important component of similarity. For the 'Version' links, it would not be surprising if being considered a version also requires the underlying model structure to be the same. Note, however, that the only two IAMs that are linked in the model family tree but fall into separate categories in Table 3.1, GCAM and Phoenix, are also connected via a 'Version' link. In this case, Phoenix is seen as a version of the ERB model (which also marks the beginning of GCAM) that is formulated specifically as a general-equilibrium model (see Table 2.3 and Appendix A).

growth models, the 545 scenarios from the Ramsey-type optimal growth models shown in Table 3.2 would still account for 46% of the entire AR5 ensemble.

Given their status as *key* differences in AR5, it is reassuring to see that most of the model dependencies identified using the method developed in Chapter 2 capture these differences. This gives additional reasons to believe that the model family tree constructed in Chapter 2 provides relevant and valid information regarding IAM relationships. Based on this, it might, however, be argued that one could simply use the information that is already contained within AR5 to identify IAM dependencies, and thus avoid the more laborious process of going through model documentation to identify model links. The method used to construct the model tree in Chapter 2, however, captures more than what the model differences listed in AR5 and the underlying frameworks do.

First, the method developed in Chapter 2 captures model dependencies that go across key differences. As already mentioned, the dependencies between GCAM and Phoenix, which are developed at the same institution by some of the same researchers, and are described as versions of the same initial model (ERB), are captured by the method developed in Chapter 2 but not by the categories in Table 3.1. Second, the method developed in Chapter 2 captures model *independencies* that are not visible from the categories in Table 3.1. IMACLIM and Phoenix, for example, are shown as independent in the model family tree because there are no links between them. Phoenix is not mentioned at all in the documentation that was collected for IMACLIM and IMACLIM is not mentioned at all in the documentation that was collected for Phoenix. In short, the model documentation indicates that IMACLIM and Phoenix have followed largely separate development trajectories (see Appendix A for more details on the model links). Separate development trajectories, all else being equal, implies a greater degree of independence in model choices and assumptions. Similarly, the model family tree shows that, while GCAM and Aim-Enduse are connected (directly) via the ERB module, and IMAGE and POLES are connected (more loosely) via the TIMER sub-model, the two branches of IAMs are not connected. This information would be lost if one simply relied on the information that is already contained in the WGIII report (shown in Table 3.1) to identify IAM dependencies. For example, it seems Phoenix and GCAM are more closely related than Phoenix and IMACLIM even though the key differences in model structure would tell us otherwise. Thus, the method developed in Chapter 2 captures both model dependencies and model independencies that would not be captured using the information that is contained in AR5.

This is not surprising when we consider the fact that model dependencies have multiple sources. Knutti and Masson (2011) and Jun et al. (2008) find dependencies between climate models developed at the same institution, between climate models sharing similar components, and between successive versions of the same climate model. All of this is likely to lead to shared model choices and assumptions, and hence, what in this thesis is referred to as model dependencies. The same is most likely true also for *IAMs* that are developed at the same institution, that use similar components, or that can be considered versions of each other. These sources of model dependencies are not necessarily aligned with key structural model differences and underlying model frameworks. Put differently, IAMs are characterised by many degrees of freedom. This means that even though the degree of foresight and the economic coverage in a model might imply a series of related model choices and assumptions, they do not dictate *all* model choices and assumptions. In short, two dimensions are not sufficient for capturing all relevant model choices and assumptions, and thus not all sources of model dependence (and independence).

Thus, the method developed in Chapter 2 provides information about relationships between IAMs that is different from the classification of IAMs provided in AR5 itself. Many IAM classifications also exist beyond the one used in the WGIII report (e.g. Dowlatabadi, 1995; Grubb et al., 2002; Hedenus et al., 2013; Löschel, 2002; Sanstad & Greening, 1998; Zhang & Folmer, 1998). Although such classifications often also point to important structural model differences, no classifications can capture all relevant sources of model dependency. In general, classifications provide a more top-down and static approach to identifying IAM similarities. They do not say anything about how IAMs have evolved and thus how some IAMs have been inspired by other IAMs. The method developed in Chapter 2 represents a different approach, which, despite limitations, captures model dependencies that are missing from existing model classifications, including the one provided in the appendix to the WGIII report.

Nonetheless, the uneven contribution from IAMs based on different model frameworks in the AR5 ensemble should be taken seriously. Given the IPCC's role in communicating the degree of certainty in the knowledge base, and given how agreement among IAMs is often perceived to represent robust insights, IPCC reports should do more to assess potential sources of IAM dependencies than what they currently do. Although, in an ideal world, all shared model choices and assumptions should be considered in the construction of IAM ensembles and in the presentation of different IAM outputs, this might not be feasible in practice. The model frameworks used, thus, appear to represent a particularly straightforward candidate. Simply listing the "key structural differences" in Chapter 6 of the WGIII report without relating any of these to the IAMs in AR5 or discussing the implications for the ensemble results is not satisfactory. Despite the fact that about half of the AR5 scenarios are generated using Ramsey-type optimal growth models, AR5 does not mention this, let alone discuss the theory and its implications.[25] This is despite a comprehensive discussion of the frameworks underpinning CBA IAMs in AR5. One reason for this might be the significant complexity of large-scale IAMs compared to CBA IAMs, whose logic and theoretical assumptions and implications have been discussed at length in the literature (e.g. Ackerman et al., 2009; Munda, 1996; Pindyck, 2013; Stern, 2013b; Weitzman, 2009). Still, given the decreasing role of CBA IAMs and the (seemingly) increasing role of large-scale IAMs,

[25] CGE models are described briefly, but not in relation to the scenario ensemble.

this appears unbalanced. The next chapter argues that there are important reasons, also beyond the goal of obtaining robust insights, to include a diversity of model frameworks in IAM ensembles.

4 The positive cost of mitigation

The previous two chapters focused on the ability to provide robust insights as a key reason for why we should care about IAM dependencies and aim to maximise diversity in IAM ensembles, in part inspired by studies that have been published in the climate modelling literature. This chapter develops a different argument for why we should care about the extent to which IAM ensembles capture the "true" uncertainty of IAM results, namely because not doing so can imply a significant risk of being wrong. This argument is based on a recently revitalised debate in the philosophy of science. The chapter starts from the observation that all the scenarios in the AR5 ensemble predict that climate mitigation will inflict economic costs in the aggregate. This leads to the impression that there must be a trade-off between climate protection and economic gains. As this chapter shows, however, this needs not be the case according to both theory and applied modelling studies. Several mechanisms that have been shown to contribute to net negative cost results in the literature can be identified. As such, the scenarios in the AR5 ensemble do not reflect the full (according to the literature) uncertainty of the cost of mitigation. This chapter argues that this uncertainty, due to the high stakes involved, is *important* and should therefore be reflected in IPCC assessment reports. Otherwise, the risk of being wrong will be high. Thus, this chapter goes beyond robustness and offers an additional reason for why we should make sure that plausible assumptions are not excluded from IPCC IAM ensembles.

In addition to this, this chapter explores potential reasons for why AR5 IAMs do not predict net negative costs of mitigation. All the cost estimates in AR5 stem from general equilibrium (optimal growth and CGE) IAMs. While some authors have argued that general equilibrium models exclude net negative costs by construction, this chapter offers a more nuanced argument. In short, although general equilibrium models can be modified to include features that enable net negative cost results, these features are difficult to implement in practice. A review of the publications that present the scenarios that are responsible for the economy-wide cost estimates in AR5 indicates that most of the IAMs include few or no mechanisms that are known to contribute to net negative cost results. Additionally, the model intercomparison studies that are responsible for the vast majority of the AR5 ensemble focused only on aspects that increase the cost of mitigation. The chapter concludes, based on this, that there is reason to believe that the AR5 IAM ensemble is biased against net negative costs. While it is beyond the scope of this chapter to investigate the process of selection and inclusion of scenarios in IPCC assessment reports, this chapter and the previous two chapters all suggest this is an important avenue for future research.

Section 4.1 presents the AR5 estimates of the cost of mitigation. Section 4.2 reviews a number of ways in which the cost of mitigation, according to the literature, might be net negative and presents a list of mechanisms that, if included in IAMs, may give rise to net negative cost results. The section also presents five examples of applied modelling studies that have predicted net negative mitigation costs. Based on this, it is argued, AR5 cost estimates do not reflect the full uncertainty of the cost of mitigation. Section 4.3 develops the argument, based on philosophy of science, for why this uncertainty is important. In short, being wrong about the cost of mitigation could have significant negative consequences. Under such circumstances, even a small possibility that the cost of mitigation might be net negative should be taken seriously. Section 4.4 shows that, although optimal growth and CGE models in principle can be modified to take into account factors that enable net negative cost results, few AR5 IAMs appear to include such mechanisms to a significant degree and all the AR5 model intercomparison studies focused on aspects that can only increase the cost of mitigation. In short, the inclusion of mechanisms that are known to contribute to net negative cost results in IAMs appears difficult and rare. Section 4.5 concludes.

4.1 AR5 cost estimates

The cost of mitigation has been central to the climate change debate since at least the beginning of the 1990s (e.g. Nordhaus, 1991b). Not surprisingly, the cost of mitigation has also been a key result in IPCC reports, presented both in the Synthesis Reports and the WGIII SPMs.

In AR5, the economy-wide costs of mitigation is measured as losses in consumption, GDP, or welfare (IPCC, 2014a). It is computed by IAMs as the difference (in the chosen metric) between a mitigation scenario and a counterfactual baseline scenario in which no (additional) climate policies are imposed. Thus, the cost of mitigation is "the difference in economic conditions relative to what would have happened without mitigation" (IPCC, 2014a, p. 448). Because partial equilibrium IAMs in AR5 do not compute economy-wide costs[26], the IAMs that are responsible for the cost of mitigation in AR5 are the ones that fall into the general equilibrium category (that is, the IAMs in the left-most column in Table 3.1).

Figure 4.1 shows the estimates of the cost of mitigation reported in the SPM of the AR5 synthesis report in the form of consumption losses. The left panel shows the increase in consumption in baseline scenarios (in 2030, 2050, and 2100) and the right panel shows the percentage reduction in consumption relative to the baseline scenarios in the mitigation scenarios (in different years and for different levels of mitigation). The figure shows that the cost of mitigation increases over time and with the stringency of the targets (i.e. costs are higher for lower CO_2eq concentrations). Scenarios that are *more likely than not* (>50-100%) to limit warming to below 1.5˚C and 2˚C reach concentration levels below 430 and

[26] Partial equilibrium models (typically) only capture energy system costs (measured as the area under the marginal abatement cost curve or energy system cost mark-up).

500 ppm CO_2eq respectively (IPCC, 2014b). Scenarios that reach concentration levels of 450 ppm CO_2eq are *likely* (66-100%) to keep warming to below 2˚C (IPCC, 2014b). The most important thing to note for the purpose of this chapter is that all the estimates of the cost of mitigation in AR5 are *positive*[27]. In addition to this, the figure also shows that the ranges of cost estimates widen over time. This is consistent with an increase in uncertainty the further into the future we go.

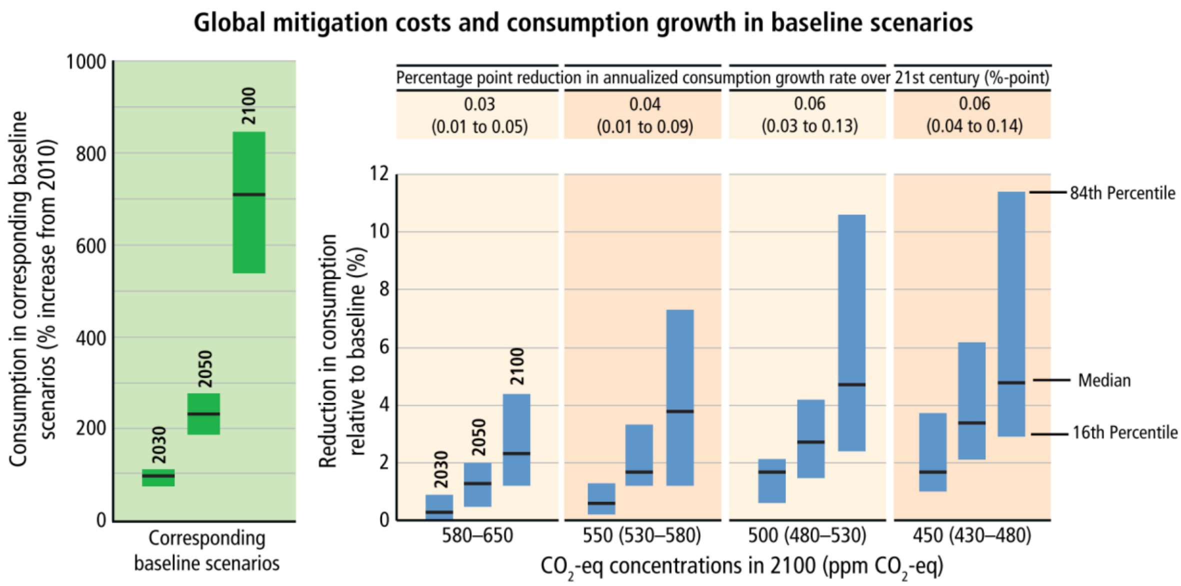


Figure 4.1 Global mitigation costs in AR5 scenarios at different GHG concentration levels in 2100. Consumptions losses (right panel) are shown relative to a baseline without climate policy (left panel). The table at the top shows annualize consumption growth reductions relative to the baseline in scenarios with different GHG concentration levels. The cost estimates do not consider the benefits of reduced climate change or co-benefits and adverse side effects of mitigation. The figure includes only results from cost-effective scenarios, i.e. scenarios that assume immediate mitigation in all countries and a single global carbon price and impose no additional limitations on technology relative to the models' default technology assumptions. Reproduced from IPCC (2014b, p. 24).

The sign of mitigation costs in AR5 implies that there is a trade-off between climate mitigation and consumption when we ignore the benefits of reduced climate change or co-benefits. This result is based on the scenarios in the AR5 ensemble and thus depends on the assumptions and model choices that are made in the IAMs that are included in this ensemble. As the next section shows, however, the literature indicates that net negative costs are possible both according to theory and according to a number of applied modelling studies. In other words, the AR5 IAM ensemble (which contains only net positive cost results) does not capture the full range of results reflected in the literature. This might lead one to question the extent to which the AR5 IAM ensemble captures the "true" uncertainty of this crucial IAM output and the reasons why it might not.

[27] Figure 6.21 in the main WGIII reports provides a more detailed picture of the cost estimates, including the full ranges (not just the interquartile ranges) and the net present values.

4.2 Negative cost – is it possible?

As noted above, the cost of mitigation reported in AR5 shown in Figure 4.1 does not include the benefits of reduced climate change, co-benefits or adverse side effects of mitigation. In that sense it only shows the "purely economic" cost of mitigation. This is in line with the majority of IAM studies. If climate benefits (i.e. avoided damages) are included, e.g. as part of consumption – the way that it typically is in CBA IAMs – the cost of mitigation can become net negative even for relatively stringent stabilization targets (depending on the model and the assumptions, particularly around discounting and the damage function, see e.g. Nordhaus (2007)). This is a crucial finding in environmental economics because it means that it is not optimal from a welfare perspective to continue with "business as usual" (and it has recently been shown that the Paris agreement passes the cost-benefit test (Glanemann et al., 2020)). The main reason why emissions reductions do not take place, according to environmental economics theory, is because climate change represents a (largely) unpriced externality. It is this kind of reasoning that led Nicholas Stern to proclaim that "climate change represents the biggest market failure the world has seen" (2013a).

In addition to this, the cost of mitigation might also be net negative if other co-benefits, such as health benefits from reduced air pollution, are included. Many co-benefits of mitigation are thought to be substantial (IPCC, 2018b). What this chapter focuses on, instead, is whether climate mitigation can be compatible with economic gains when we ignore the benefits of reduced climate change and other co-benefits. That is, *can climate policy be "good" for the economy?* As this section shows, this is possible both according to theory and according to applied modelling studies.

4.2.1 According to the literature, yes

A number of academic debates in the 1990s revolved around the question of whether environmental policy could be good for the economy. The 'Porter hypothesis' (Porter & Van Der Linde, 1995) asserted that firms can benefit from environmental regulation because this can spur innovation, which in turn can increase productivity. On an aggregate level this could mean that there is no trade-off between economic growth and environmental protection, but instead a win-win situation. Similarly, the so-called 'no regret' emissions reductions potential, which was closely linked to the 'efficiency gap controversy', was also heavily debated in the 1990s (Hourcade, 1993). The 'no regret' potential refers to an amount of emissions that can be reduced at net negative costs (Maréchal, 2007). The most famous example of such a potential is the investment in energy efficiency, which, because it is often not undertaken in real life, has led to the notion of an ‘efficiency gap’ or an ‘efficiency paradox’ (Marechal & Lazaric, 2010)[28]. Related to how climate policies are implemented in practice, the ‘double dividend’ hypothesis suggested that increased environmental taxes could improve not only environmental conditions, but increase economic efficiency if combined with the reduction of other (distorting) taxes (Goulder, 1995). This is also known as the ‘revenue-recycling effect’.

What these concepts have in common – besides being related to the possibility of net negative mitigation costs – is that empirical evidence has neither confirmed nor denied their validity. Questions remain regarding the ability to reap the associated economic benefits in the real world. Ambec et al. (2013) conclude, almost 20 years after the Porter hypothesis was conceived of, that the empirical evidence is mixed. In some cases, environmental regulation appears to have improved business performance, in other cases not. When it comes to the efficiency gap, economists and energy system analysts still debate whether it exists and what the reasons behind it might be (Allcott & Greenstone, 2012; Gerarden et al., 2017; Jaffe & Stavins, 1994) and, related, the real costs of no regret potentials (Marechal & Lazaric, 2010). Similarly, although the choice of revenue recycling is held to be a key determinant of the cost of mitigation (Barker et al., 2006; Bye et al., 2002; Repetto & Austin, 1997) – and even though authors have found evidence of strong double dividends (Bosquet, 2000; Bye et al., 2002) – researchers are still not in agreement as to whether revenue recycling can lead to net negative costs overall. As was already the case in 1995, the weak version of the hypothesis – revenue recycling through cuts in distortionary taxes leads to cost reductions relative to the case where revenues are returned lump-sum – is widely supported, but the empirical evidence for the strong version – replacing distortionary taxes with environmental taxes lead to zero or negative costs – is mixed (Goulder, 1995). According to Guivarch et al. “the answer [regarding the potential of revenue recycling] has not really been resolved” (2011, p. 2).

Thus, if we look to the literature (on the energy efficiency gap, the Porter hypothesis, the double dividend, and revenue recycling) the answer to the question of whether net negative cost of climate mitigation (and environmental protection more generally) is possible appears to be “it depends”. More recently, several authors have again explored the possibility of net negative costs in relation to the concept of Green Growth, which was popularised by institutions such as the World Bank, OECD, and UNEP around 2011-2012 (OECD, 2011; UNEP, 2011; World Bank, 2012) [29]. If we combine the above

[28] The existence of a no regret potential does not on its own imply that emissions reductions at the level needed to meet the Paris target can be undertaken at net negative costs, but it will reduce overall costs.

[29] There are two versions of the green growth concept. The strong version asserts that environmental protection can positively promote economic growth and the weak version simply asserts that proper environmental protection

concepts with the ideas that have been discussed in relation to the green growth debate, we arrive at a number of mechanisms that, if they are present, might give rise to, or at least contribute to, net negative mitigation costs. These mechanisms are listed in Table 4.1 and discussed below.

Table 4.1 Mechanisms that might enable, or at least contribute to, net negative mitigation costs.

Green Growth Mechanisms	
Keynesian green stimulus ('Green New Deal')	Climate policy can act as a stimulus on the economy, which will increase employment and capital utilization, and thus spur growth.
Limits to brown growth	Increasing environmental damages will sooner or later hamper economic growth itself. Climate policy can move us in the direction of a better long-term growth path.
Market failure corrections	Climate policy can be used to correct existing market failures and thus increase the overall efficiency of the economy.
Revenue recycling	Climate taxes can be used to decrease other distortionary taxes.
No-regret options	Climate policy can be used to incentivise cost-effective choices that are not currently made (e.g. investments in energy efficiency).
Knowledge spillovers (R&D)	Climate policy can be used to increase economic productivity by remedying current underinvestment in low-carbon R&D caused by innovation externalities.
Under-employed production factors	Climate policy can help increase the employment of currently under-employed production factors (labour and capital).
Learning-by-doing	Climate policy can be used to reduce the future cost of energy (and thus increase economic productivity) by remedying current underinvestment in low-carbon technologies caused by learning externalities.
Schumpeterian green revolution	Climate policy can unleash a wave of innovation that will ultimately transform the economy and bring about a "green industrial revolution".

can be compatible with economic growth (i.e. environmental protection does not necessitate a reduction in economic growth) (Jacobs, 2013).

Keynesian green stimulus ('Green New Deal')

The original case for green growth in the aftermath of the 2008 financial crisis was based largely on the Keynesian argument that an economy in recession can be stimulated back into growth through environmental measures (e.g. Jones (2009), New Economics Foundation (2008)). The argument, also known as a 'Green New Deal', rests on the idea that in a recession, governments should stimulate demand by replacing lost private demand with public expenditure (Jacobs, 2013). This expenditure need not be green for it to have the desired effect, but if green investment is going to be needed in the near future regardless, it represents a particularly good choice. Investment in green infrastructure represents one such option (Hallegatte et al., 2012). Many have also argued that since environmental measures are often labour intensive, they offer better economic growth per dollar spent than other measures (Jacobs, 2013; Mark et al., 2011).

Critics of the Keynesian stimulus idea argue that the effects are only short term. According to Hallegatte et al. "demand-led and Keynesian effects" become important "when actual production is more constrained by demand than by production capacity (i.e., in situations of high unemployment and low utilization of production capital)" (2012, p. 8). While many proponents of Keynesian policies will agree that stimulus measures are only meant to be short-term, some have, however, argued that environmental measures can also drive economic growth in the medium- and long-term (e.g. Spencer et al. (2012)).

Limits to brown growth

Another argument says that the current pattern of economic growth, if continued into the future, will ultimately be bad for economic growth itself. This argument is based on the idea that environmental damages, when they get bad enough, will start to also damage productivity. Current patterns of growth, the argument goes, fail to take into account the increasingly negative impacts of increasing climate damages on natural capital and labour (Hallegatte et al., 2012). That is, long-term economic output can be increased by climate policy because it will reduce damages to physical (e.g. from extreme weather events) and human (e.g. from air pollution) capital, both of which affect productivity. This view can be contrasted with the traditional view that environmental policies will have a negative impact on capital because it constrains available technologies and forces polluting technologies into early retirement.

Jacobs (2013) notes, however, that even though many argue that sustainable management of resources might generate growth, many argue that unsustainable exploitation will generate more growth. The latter is, after all, how developed nations grew: by transforming natural capital into physical capital, which led to higher productivity and ultimately higher economic output. The answer will ultimately depend on one's assessment of the current and future state of the environment and its impact on the economy. While proponents of green growth will argue that although "brown growth" was possible in the past, we have now reached a point at which the environment has become so scarce that this is no longer viable, proponents of continued “brown growth” will argue that this time has not yet arrived.

Market failure corrections

A wide range of market failures (beyond the climate change externality) are known to be present in the economy. If climate policy is used to correct some of these market failures, the net result could be an increase in economic efficiency rather than a decrease. Many of the concepts that were discussed in the 1990s fall into this category.

Revenue recycling (double dividend)

Environmental taxes can be used to correct the economic inefficiency of the current taxation system. ‘Revenue recycling’, if done the right way, can thus lead to a ‘double dividend’. This possibility is also referred to as ‘environmental tax reform’.

No-regret options

‘No-regret’ options imply that (at least some) emissions reductions are possible at net negative costs. Investment in energy efficiency, for example, is widely thought to be good both for the environment and the economy[30].

Knowledge spillovers (R&D)

Invention and innovation failures due to underinvestment in basic research and development (R&D) are known to be caused by knowledge spillovers (Ackerman et al., 2010; Edenhofer et al., 2013; Nemet, 2013). Given the crucial role of innovation for productivity and growth, correcting the effects of knowledge spillovers in the case of low-carbon technologies might benefit both the climate and the economy. Although not exactly the same, this idea is also related to the ‘Porter hypothesis’ discussed above (or a macro-scale version of this hypothesis (Hallegatte et al., 2012)). Policies that can be used include direct government investments in R&D, innovation prizes, patenting systems, and carefully designed tax breaks and subsidies (Stern, 2014). Hallegatte et al. (2012), however, also note that green investments might sometimes crowd out more profitable investments.

Learning by doing

The basic idea behind learning-by-doing, which dates back to Arrow (1962), is that the cost of producing a good declines with the cumulative production of the good. Learning-by-doing is a cause of

[30] If energy efficiency leads to increased economic output this also leads to a rebound effect, which means that the environmental benefits are smaller than if there were no economic gains.

market failure if firms cannot capture the cost reductions caused by their own production of the good. For renewable energy technology in particular, there is ample evidence that the cost has declined with cumulative installed capacity (Rubin et al., 2015; Samadi, 2018).

In an extensive meta-analysis of cost estimates Barker et al. (2006) found that the treatment of technological change is a major determinants of the cost of mitigation. Together, the effects of R&D and learning-by-doing on technological change are often captured using the term *induced technological change* (ITC). In a large model intercomparison study of 10 IAMs, Edenhofer et al. (2006) showed that ITC reduced abatement costs in all models. Bosquet (2000) also found that technological change increases the economic benefits of environmental tax reforms.

Under-employed production factors

Production factors, including capital and labour, might be under-employed not only in a recession, but also for structural reasons (Hallegatte et al., 2012). Climate policy might help increase the employment of several production factors. Both Bosquet (2000) and Bye et al. (2002) find that the presence of structural unemployment also increase the economic benefits of revenue recycling.

Schumpeterian green revolution

Certain authors have argued that environmental policy can serve as a new engine of growth for the global economy (Perez, 2009, 2017; Stern, 2014). According to Perez, policies that take environmental threats as opportunities and combine smart green growth with full global development can unleash a "global sustainable Golden Age" (2017, p. 3). According to Stern, climate policy can be used to bring about "a new green industrial revolution" (2009). While invention and innovation is also at the core of this argument, it extends beyond the more narrow confines of knowledge spillovers (discussed above) and borrows from Schumpeterian theories (presented e.g. in Schumpeter (1983)) to make more general claims regarding economic development. The idea is that sufficient levels of environmental policy will unleash a wave of innovation that will ultimately transform the economy.

The above mechanisms, summarized in Table 4.1, can give rise to, or at least contribute to, net negative mitigation costs. They are not mutually exclusive. Learning-by-doing and knowledge spillovers, for example, which are listed under market failure corrections, are also central to the idea of a Schumpeterian green revolution. Limits to "brown growth", due to increasing environmental degradation, could also be considered an (inter-temporal) market failure. This is indeed similar to how the climate change problem is conceived of in cost-benefit IAMs, the main difference being that here we are not talking about utility or welfare, but economic output.

The ultimate impact of climate policy on the economy will depend on the balance of mechanisms pulling in different directions. Can a Keynesian stimulus have a long-term effect on economic output, for example, or is this only a short-term fix? Will the increasing environmental damages associated with the current growth path outweigh the benefits of staying on this path? Are the constraints imposed by climate policy on capital and labour larger or smaller than the benefits in terms of innovation and increased energy efficiency? What effect dominates, knowledge spillovers or crowding out? Overall, *how large are existing market failures and to what extent can climate policy correct them?* Based on the literature, we do not currently have a clear answer to these questions. Expert opinions differ, often considerably. The fact that no data exists on the impacts of climate mitigation on the scale required to meet the Paris target does not help resolve the issue. While we certainly cannot conclude that climate policy will generate net negative costs, the main takeaway from this section is that we also cannot conclude that climate policy *cannot* generate net negative costs. The latter, however, is what is indicated by AR5. The next section presents a handful of applied modelling studies that further corroborates the possibility of net negative cost results.

4.2.2 According to applied modelling studies, yes

This section presents five examples of models that have predicted net negative mitigation costs. While no attempt has been made to offer a complete list of models that have produced such results, an effort has been made to capture a variety of different types of models (including two CGE, one optimal growth, one macroeconometric, and one systems dynamics model). Several of the examples are taken from the review of model structures for win-win strategies by Wolf et al. (2016). Because the main point in this section is to show that such model examples *exist*, the discussion of each model will be brief. Some of the model assumptions will, however, be discussed in more detail in section 4.4.

T21

The UNEP Green Economy Report (UNEP, 2011) uses a macroeconomic model based on systems dynamics called Threshold 21 (T21) to show that a green economy can grow faster than a 'brown' (baseline) economy. T21 was developed by the Millennium Institute[31] in the 1990s (Qu et al., 1998) to study strategies for medium to long-term development and poverty reduction. It combines optimisation in the energy sector with econometrics in the economic sectors. A key feature of T21 is the fact that it takes into account the role of natural resources in production (in addition to capital and labour). According to the UNEP Green Economy Report, "the inclusion of natural resources as a factor of production distinguishes T21 from all other global macroeconomic models" (UNEP, 2011, p. 24). While economic output is higher in the baseline scenario than in the green scenarios for T21in the first

[31] https://www.millennium-institute.org/

7-9 years, economic growth in the green scenarios overtake growth in the baseline scenario and lead to permanently higher economic output after this initial period of green investment.

E3MG

The macro-econometric model E3MG (now called E3ME) has generated scenarios that predict economic benefits from climate policies on several occasions (e.g. Barker et al., 2006; Barker & Şerban Scrieciu, 2010). In fact, results from E3MG depicting net negative costs from climate mitigation was included in the IPCC AR4 scenario ensemble (see Figure 11.17 in IPCC (2007)). The distinguishing features of E3MG are, according to Barker and Scrieciu (2010), its Keynesian demand-led approach, its representation of the economy without assuming equilibrium, and its assessment of policies in a non-optimal environment. E3MG assumes increasing returns to production and under-employment of labour source in the baseline (Edenhofer et al., 2006). This means that climate policy can lead to an increase in investment which in turn leads to an increase in economic output. These features allow positive macroeconomic effects of climate mitigation, in both the short and long terms, although both the magnitude and sign of economic impacts depend on the specifics of how policies are implemented (including how carbon revenues are used). E3MG has even predicted that economic output can increase with the stringency of climate targets (Barker & Şerban Scrieciu, 2010).

GEM-E3

The CGE model GEM-E3 (EU Science Hub, 2019) computed positive economic impacts of climate policy in the study *A New Growth Path for Europe* by Jaeger et al. (2011). The version of GEM-E3 used in this study was enhanced specifically to include key features such as learning-by-doing and to allow expectations to influence economic dynamics. The natural rate of unemployment was also allowed to vary. Together, these elements allowed GEM-E3 to model a virtuous cycle of feedback between investment into green technology, learning by doing, and investor expectations. As a result, the version of GEM-E3 used in this study was able to depict a transition from what was seen at the time to be an inferior equilibrium (with high unemployment and low growth) to a superior equilibrium (with lower unemployment and higher growth) in which GHG emissions were also reduced (Jaeger et al., 2011).

IMACLIM

A follow-up study by Jaeger et al. (2015) also included the CGE model IMACLIM (in addition to GEM-E3), which was used to compute scenarios in which green policies coordinated investor behaviour in a way that stimulated the economy and thus increased economic growth relative to the baseline. The version of IMACLIM used in this study also included learning by doing as well as economic frictions. These features, which are default features of IMACLIM (Sassi et al., 2010; Waisman et al., 2012), were, however, not sufficient to generate net negative costs on their own: introducing a carbon price to the default version of IMACLIM in this study led to a reduction in GDP. In order to depict pathways with net economic benefits, two additional assumptions were therefore imposed: First, revenues from carbon taxes were recycled, and second, a specific stimulus package that redirected finance from high to low carbon activities was imposed. The result of all these assumptions was a significant increases in GDP relative to the baseline at all time periods in the climate policy scenario (Jaeger et al., 2015).

FEEM-RICE-FAST

A version of the optimal growth model FEEM-RICE, called FEEM-RICE-FAST (Bosetti, Carraro, & Galeotti, 2006), predicted GDP gains from emissions reductions in the Innovation Modeling Comparison Project (IMCP) (Edenhofer et al., 2006). The baseline in FEEM-RICE-FAST assumes market imperfections due to externalities in R&D investment: Regions under-invest in R&D due to assumed non-cooperative behaviour (Edenhofer et al., 2006). When climate policy is introduced, regions are induced to increase their R&D investments, bringing these closer to the cooperative, optimal, levels. The net negative cost results, however, also hinges on two more assumptions, namely the existence of learning-by-doing and limited crowding out. If the learning-rate is slow and the crowding out effect is large, costs are no longer net negative (Edenhofer et al., 2006).

All of the above modelling studies represent serious attempts to depict real-world possibilities. Although some of the assumptions underlying some of the model runs might be thought of as optimistic[32], none of them are deemed implausible (at least by the authors themselves).

Based on theoretical arguments and applied modelling studies found in the climate policy literature – of which this section has provided only a snapshot – it appears that even the *sign* of the cost of mitigation is uncertain. The possibility that mitigation costs can be net negative, however, is not captured by AR5 results[33]. The next section argues that this possibility is important and that it should therefore be recognised in IPCC reports.

[32] According to Edenhofer et al, “in the case of FEEM-RICE-FAST the negative costs are the consequence of the optimistic assumptions on the effects of R&D investments and of the role that stabilization targets have in inducing more R&D investments” (2006, p. 76).

[33] It should be noted that two of the five models discussed, GEM-E3 and IMACLIM, are included in the AR5 IAM ensemble. The scenarios generated by these two models in the AR5 ensemble do not, however, predict net negative costs (but, in the case of IMACLIM, unusually high positive costs). The potential reasons for this will be discussed in section 4.4.4.

4.3 When uncertainty matters

A question asked by philosophers of science in the 1950s was whether scientists, when giving advice to decision makers, should consider the potential consequences of error when deciding what to say (Rudner, 1953). In a well-known article and book on values in science, Heather Douglas (2000, 2009) recently revised this debate[34].

The premise of Douglas' argument is that all humans have a general moral responsibility to consider the consequences of error when making choices. It is widely accepted that people are responsible for side effects of their actions when these are caused by reckless or negligent behaviour. Being *reckless* means that one is aware of the risks caused by a choice, and that these risks are unjustified, but one still goes ahead. An example of this is speeding on city streets for the fun of it (as opposed to speeding in order to get a seriously injured person to the hospital). Being *negligent* means that one does not bother to properly evaluate obvious risks or fails to think about potential consequences. An example of this is making a bonfire on a dry and windy summer day with no thought as to how to control the fire. The argument in this case is that the person *should* have foreseen the potential problems and planned accordingly. If the fire is made with clear awareness of the risk, but the person doesn't care, then the action would be reckless. In Douglas' own words "recklessness is proceeding in the face of unreasonable risk; negligence is the failure to foresee and mitigate such risk" (Douglas, 2009, p. 70).

Douglas' main point is that one can be negligent or reckless not only in actions, but also in making descriptive or empirical claims. According to Douglas, "making empirical claims should be considered as a kind of action, with often identifiable consequences to be considered" (Douglas, 2009, p. 70). As an example, she considers the question of whether one should report an unattended briefcase. Whatever the choice, there are clear risks of error. If one reports the briefcase, and it is not a bomb, it will disrupt people's daily lives and take away resources from more important tasks. If one does not report it and it is a bomb, serious harm and death might result. Clearly, the latter result is much worse. But probabilities should also be considered[35]. In the same manner, Douglas argues, *scientists* are also morally responsible for the consequences of making inaccurate or unreliable empirical claims.

This argument is not universally accepted and some have argued that scientists should be exempt from this moral responsibility (e.g. Lübbe, 1986). This argument is generally based on the view that such

[34] See Elliot (2017) and Kincaid et al. (2007) for more on this debate.

[35] The probability will depend on contextual factors, such as where the briefcase is spotted. If spotted in a classroom known to be used by a particularly absentminded colleague, the probability that the briefcase is *not* a bomb (but the colleague's briefcase) appears to be much higher than if the briefcase is spotted on a busy subway station.

considerations would impose too much of a handicap on scientists' work. The strong version of the argument says that it would impose unreasonable limits on scientific freedom, which is essential to the pursuit of knowledge. Because this knowledge is so valuable to society, it justifies the exemption of scientists from such moral obligations. Against this, Douglas points to the fact that we already do not allow the pursuit of knowledge to trump all other values. The value of knowledge is always weighed against other values. For instance, we place limits on the use of research subjects. In other words "the knowledge produced by scientists should not be and is not considered priceless" (Douglas, 2009, p. 77). The weaker form of the argument says that, even if we don't consider science to be *priceless,* requiring scientists to consider the consequences of their work still places a *too large* burden on the scientists. Against this, Douglas argues, "the price of morally exempting scientists from the general responsibility to consider the consequences of errors looks much higher than the price of having scientists shoulder this burden" (Douglas, 2009, p. 78). Given cases where moral exemption would be very harmful and the absence of clear boundaries of cases in which it would and wouldn't, Douglas argues, blanket exemption is not defensible. Thus, Douglas rejects the argument that scientists should be exempt from the moral responsibility of making inaccurate or unreliable claims.

This does not imply that scientists are responsible for every use or misuse of their work, which would be an unreasonable expectation. Just as humans should only be held responsible for the unintended consequences of their actions under certain conditions (e.g. when they are reckless or negligent), the same should be true for scientists. That is, scientists – like other people – are only responsible for what is reasonably foreseeable. Douglas uses the discovery of the neutron in nuclear physics by James Chadwick and the research that followed in the 1930s as an example. This discovery later led to the development of the atomic bomb. However, it was not until the discovery of fission in December 1939 that the atomic bomb could be conceived of. Thus, we could not have expected Chadwick to foresee this possibility, but we could expect all nuclear physicists working in this area after 1939 to foresee this. In short, "the moral burdens on scientists are not unlimited. They are held to only what can be foreseen, and thus discussed and considered" (Douglas, 2009, p. 84).

It is in light of this that Douglas (2009), in line with Rudner (1953), argues that scientists should consider the consequences of being wrong when deciding whether or not to make a claim. Rudner (1953) argues that scientists should weigh *the importance of the uncertainty* based on *the consequences of making an error*:

> "since no scientific hypothesis is ever completely verified, in accepting a hypothesis the scientist must make the decision that the evidence is sufficiently strong or that the probability is sufficiently high to warrant the acceptance of the hypothesis. Obviously our decision regarding the evidence and respecting how strong is "strong enough", is going to be a function of the importance, in the typically ethical sense,

> of making a mistake in accepting or rejecting the hypothesis. Thus, to take a crude but easily manageable example, if the hypothesis under consideration were to the effect that a toxic ingredient of a drug was not present in lethal quantity, we would require a relatively high degree of confirmation or confidence before accepting the hypothesis - for the consequences of making a mistake here are exceedingly grave by our moral standards. On the other hand, if say, our hypothesis stated that, on the basis of a sample, a certain lot of machine stamped belt buckles was not defective, the degree of confidence we should require would be relatively not so high. How sure we need to be before we accept a hypothesis will depend on how serious a mistake would be" (Rudner, 1953, p. 2 italics in original).

Douglas (2009), similarly, uses the example of a scientist that discovers a correlation between a particular pollutant and respiratory deaths to explain how scientists should exercise their moral responsibility. If the pollutant in question is cheap and easy to control, the consequences of incorrectly accepting the claim as reliable appears much less severe than the consequences of incorrectly rejecting the claim as reliable. In this situation, Douglas argues, the scientist should note the uncertainty, but suggest that the evidence sufficiently supports the claim. When scientists are considering potentially catastrophic consequences of making an error, such as was the case when scientists considered the potential effects of the first atomic bomb test in New Mexico in 1945, less uncertainty is tolerable for a claim that the testing is safe.

The key point for both Rudner (1953) and Douglas (2009), thus, is that the *importance* of the uncertainty – and thus the amount of uncertainty that scientists should tolerate in making claims – depends on the consequences of being wrong. When the consequences of being wrong are significant, scientists should require a greater level of certainty when making a claim. When consequences are relatively insignificant, we need not worry so much about being wrong. In short, the higher the stakes, the more important the uncertainty.

4.3.1 Implications for IAM research

If we accept Douglas and Rudner's argument, researchers are morally responsible for the consequences of making inaccurate or unreliable claims also based on IAMs. As has already been noted, the cost of mitigation is a central IAM output. So far, this chapter has presented theoretical arguments and provided examples of applied modelling studies that suggest mitigation costs *could be* net negative. The AR5 IAM scenarios, however, do not reflect this possibility. In order to assess the importance of the uncertainty associated with the cost of mitigation, we now consider the consequences of being wrong regarding this IAM result.

In order to consider the consequences of being wrong, we assume that the “true” cost of mitigation is net negative. We then ask what the consequences are of incorrectly predicting positive costs.[36] If the cost of mitigation in reality is net negative, this means that we are currently on a sub-optimal (from a purely economic point of view) pathway. The consequences of not yet having reduced global GHG emissions amounts to decades of lost economic benefits. In the counterfactual (optimal) pathway, one or several of the mechanisms listed in Table 4.1 (reduced damages to capital and labour from climate change, corrections of existing market failures, economic stimulus through periods of slow growth, and possibly even a new wave of green innovation) could have led to a win-win situation compared to where we are at. Given the global magnitude of the economic effects and the fact that three decades have passed since the first IPCC report, if the “true” cost of mitigation is (and has been) net negative, the economic consequences of not having mitigated are most likely considerable. If we include the environmental consequences, the cost is larger. This is because the counterfactual (optimal) pathway implies not only economic benefits, but environmental benefits: In this pathway, more mitigation would take place[37]. As evidenced by the IPCC’s SR15 (IPCC, 2018a), even a 0.5˚C difference in global warming can have enormous consequences. In the worst case, the difference between the pathway that we are currently on and the optimal from an economic point of view (corresponding to the case in which the cost of mitigation is net negative) could amount to the difference between a stable climate and a climate in which irreversible tipping points are reached.

Thus, if we assume the “true” cost of mitigation is net negative, the difference between the sub-optimal pathway that we are currently on and the optimal pathway that we could have been on, is likely large both in economic and environmental terms. There is one last question, however, and that is how large the impacts of IPCC cost estimates are. That is, how much do these results affect climate policy and action? If the influence of IAM cost results are small, the consequences of being wrong would also be small. In this case, the uncertainty of the results would not be important because the results are not important (in the sense of having an impact on climate policy and thus mitigation). There could be two reasons why IAM cost results don’t have a material impact on climate policy: either policymakers don’t take IAM cost results seriously or the aggregate cost of mitigation is of relatively minor importance to

[36] Given the uncertainty and variability of IAM results, we here equate “being wrong” about the sign of the cost of mitigation from the perspective of the IPCC reports as a situation in which all IAMs in the ensemble predict net positive costs when the true cost is net negative. IAM ensembles can of course also be wrong about the magnitude of costs, whether these are negative or positive. But because of the significance of the sign, and in order to simplify the discussion, we here focus only on the sign of the cost.

[37] Note that the optimal pathway when the cost of mitigation is net negative does not necessarily imply an instantaneous reduction of emissions to zero. Even in the case where the cost of mitigation is net negative, a smooth reduction of emissions would likely be preferred to a sudden reduction.

policymakers. The first explanation leads us into a somewhat paradoxical situation: if IAMs are designed to inform climate policy, but IAM results don't have an impact on policy, then either IAM research is not worthwhile, or policymaking should be changed such that it takes IAM results more seriously. If one concludes the former, that IAM research is not worthwhile, then IAM researchers have bigger problems to worry about than the risk of producing inaccurate or unreliable results. In this situation, although it would be true, it would also be self-deprecating for IAM researchers to argue that they don't have to worry about the risk of being wrong because their results don't matter. If one instead concludes the latter, that policymakers should take IAM results seriously, then one should also care about the risk of being wrong regarding the cost of mitigation.

The second explanation is likely to hold true at least to some extent: policymakers care about much more than the aggregate global cost of mitigation. They are likely to care more about domestic gross costs, and care about energy security and affordability, not just aggregate costs. In addition to this, policies interact and climate policies, like any other policies, are rarely "only" climate policies: policymakers are constantly weighing multiple goals against each other and a variety of constraints. Nonetheless, costs and economic growth are and have been key concerns for policymakers in the last decades. Costs were, among other things, one of the main arguments used by the US when leaving the Kyoto agreement. Even though IAMs are far from the only input to climate policy, they still influence the public debate and inform climate policy both nationally and internationally (Krey et al., 2019a). Although we do not know the magnitude of the impact of IAM cost estimates on climate policy, we can be fairly certain of the direction of the impact.

In both cases, IAM researchers should act as if the results of their work are taken seriously and have an impact on climate policy – informing climate policy is, after all, the main purpose of IAMs. Thus, assuming that IAM research is worthwhile and that IAM results either already have an impact[38] or that IAM researchers want it to have an impact on climate policy[39], we can consider how IAM cost estimates might affect climate policy, and thus what the consequences of being wrong regarding this IAM result might be.

---

[38] This does not just include direct impact. The cost of mitigation is also part of the public debate around climate change, and this in turn has a significant impact on climate policy.

[39] This does not imply that IAM results will be fed into policymaking in a simple linear manner, nor that their results will not be challenged in the policymaking process. Policymakers are not passive consumers of scientific research. Nor do IAMs, in the vast majority of cases, provide clear recommendations for policy. Ultimately, politics entails evaluating different sources of information and weighing up multiple considerations against each other. The interplay between science and policy is messy (see e.g. Owens (2015)).

In assessing the consequences of being wrong, it is again helpful to consider the counterfactual: assuming the true cost of mitigation is net negative, what might have happened if IAM ensembles (correctly) predicted this?[40]

First, nations most likely would have been more willing to sign up to an international climate agreement if IAM research indicated that mitigation presented a win-win option. The cost of mitigation was reported in IPCC reports for the first time in the second assessment report (AR2) in 1995 (IPCC, 1995). This was around the time when the negotiations leading up to the Kyoto agreement were taking place. Several authors have, according to Bowen and Fankhauser, "long argued that a key barrier to reaching an international agreement on climate change is the burden-sharing focus of the UN Framework Convention on Climate Change" (Bowen & Fankhauser, 2011). The belief that mitigation would be costly most likely contributed to this. In contrast to this, the concept of win-win policies that benefit not only the climate but also the economy "tries to make environmental policies easier to implement in spite of political obstacles, and to increase the social and political acceptability of environmental policies" (Hallegatte et al., 2012, p. 30). Among other things, economic cost was the main argument used by the US when leaving the Kyoto agreement. If IPCC reports had instead indicated that costs would most likely be net negative, it seems plausible that the US would have been less likely to leave the agreement or at least found it harder to justify doing so. If the US had stayed in the Kyoto agreement, this would have also most likely had a positive impact on global mitigation efforts. The US was not only the biggest emitter at the time (a role which has now been taken by China[41]), but also a world leader in many other respects (e.g. in terms of technology).

Second, international climate agreements would have likely been more ambitious. The 2˚C target that was finally agreed in Paris in 2015 was the result of a long and tedious process stretching over many years. Costs have always been an important consideration for the parties to the agreement. If IPCC reports had indicated all along that costs would most likely be net negative, the incentives for stronger action would have also been stronger.

Third, net negative cost results would have most likely made it easier for policymakers also at the domestic level to implement climate policies. This is because climate policies would have been seen as less of a threat to economic development, which remains a primary goal.

---

[40] If IAM ensemble results are "wrong" about the cost being net negative when they predict only net positive results, they can be said to be "correct" if they predict only net negative cost results (or at the very least, if they predict net negative costs on average).

[41] In absolute levels, not per capita (Friedlingstein et al., 2019).

It should be noted that even if the true cost of mitigation is net negative this does not imply that there are no costs associated with climate mitigation. Net negative costs do not prevent gross costs from being large. These gross costs will be felt to different degrees by different industries, regions, and people. The transformation of the energy system will have large negative consequences for particular industries, notably the fossil fuel industry. The transformation also implies significant risks of stranded fossil fuel assets, which again has macroeconomic and distributional consequences (Mercure, Pollitt, Viñuales, et al., 2018)[42]. And on a domestic level, the implementation of climate policy will impose certain costs for certain parts of the population. The gilet jaunes (yellow jacket) movement in France provides a good illustration of how people might react to these very real costs. These gross costs, which, if not implemented alongside counterbalancing redistributive policies, will be felt unequally by people and industries. Thus, such costs will most certainly still represent a barrier to the implementation of climate policy, even if the net global cost of mitigation is negative – and even if IAM results reflected that. Nonetheless, net negative cost results might have shifted the debate from a question of overall economic output and burden sharing to a question of how to distribute the gains in a just manner.

It is impossible to say exactly how much more mitigation would have taken place if IAM ensembles (correctly) predicted that the cost of mitigation would be net negative. More research into the use of IAM results and the effects of IPCC results on climate policy and action is needed. But, even then, it would be impossible to say exactly what would have happened if things had been different. Nonetheless, even though the magnitude of the impacts on climate policy of the cost of mitigation is unknown, the direction seems clear: net negative cost results would not have led to *less* climate mitigation. We can assume, with a high degree of confidence, that if IAMs predicted net negative costs all along, earlier and stronger mitigation would have resulted. Thus, although we cannot know the magnitude of this effect, we can argue that the consequences of IAMs being wrong about the cost of mitigation (assuming the cost of mitigation in reality is net negative) includes missed economic opportunities and increased climate damages. Given the large difference in climate damages caused by 1.5°C of global warming and 2°C of global warming, and how close we are to not being able to make either of these targets today (IPCC, 2018a), the consequences of being wrong about the cost of mitigation might, in the worst case, amount to irreversible environmental damages.

In summary, being wrong about the cost of mitigation is likely to have negative consequences. The scale and nature of the climate change problem implies that the lost economic benefits and increased

[42] I am a co-author of this paper.

climate damages are likely to be substantial. Assuming IPCC assessment report results have a material impact on mitigation, the consequence of being wrong is significant.

This chapter has argued that the uncertainty regarding the sign of the cost of mitigation is considerable. The potential magnitude of the consequences of being wrong implies that this uncertainty is also *important*. The AR5 IAM ensemble, however, does not reflect this uncertainty.

The argument that AR5 leaves out *important* uncertainties regarding the cost of mitigation concludes the first part of this chapter. In order to link this conclusion to the arguments and findings in Chapters 2 and 3, and in order to present a more complete assessment of the AR5 cost estimates, the rest of this chapter examines potential reasons for why the AR5 IAM ensemble does not capture the possibility of net negative mitigation costs. As in previous chapters, given the large number of IAMs and the complexity of each IAM, the analysis will not be based on a detailed investigation of all the assumptions that determine cost estimates in the individual IAMs that in AR5 produce economy-wide cost estimates. This means that it is difficult to draw definite conclusions. The analysis nonetheless provides a relatively thorough assessment of the potential reasons why AR5 IAMs do not produce net negative cost results based on a review of model publications, model documentation, the WGIII report, and more general arguments drawn from the economics literature.

4.4 Potential reasons why net negative costs do not appear in AR5

4.4.1 Because general equilibrium models exclude the possibility by construction

The first thing to note is that all the economy-wide cost estimates in AR5 stem from general equilibrium IAMs. There are 20 general equilibrium IAMs in AR5 (shown in Table 3.1). As shown in Chapter 3, eight of the 14 most influential IAMs in AR5 are general equilibrium IAMs. These eight models are either optimal growth or CGE models (Table 3.2). A closer look at the remaining 12 general equilibrium IAMs in AR5 (shown in grey in Table 3.1) reveals that these are also either optimal growth or CGE models. By looking at the number of scenarios generated by each model, we find that 79% of the scenarios generated by general equilibrium IAMs in AR5 stem from optimal growth models and that 21% stem from CGE models. Table 4.2 lists the general equilibrium IAMs in AR5 and the versions used in AR5, and the number of scenarios generated per IAM.

Table 4.2 General Equilibrium IAMs in AR5. Source: WGIII AR5 Scenario Database (IAMC, 2014).

	Model types, names, and versions included in AR5		**# of scenarios in AR5**
	Optimal Growth	**CGE**	
Perfect foresight IAMs			
	REMIND (1.1, 1.2, 1.3, 1.4, 1.5)		158
	MESSAGE* (V.1, V.2, V.3, V.4)		140
	WITCH (AME, AMPERE, EMF22, EMF27, LIMITS, RECIPE, ROSE)		132
	MERGE-ETL (2011)		48
	MERGE (AME, EMF22, EMF27)		44
	BET (1.5)		23
	EC-IAM 2012		21
	GRAPE (ver1998, ver2011)		14
	MARIA23_org		5
		iPETS (1.2.0)	4
Myopic (recursive dynamic) IAMs			
		IMACLIM (v1.1)	53
		Phoenix (2012.4)	31
		ENV-Linkages (WEO2012)	17
		FARM (3.0)	12
		GEM-E3-ICCS	11
		WorldScan2	8
		SGM	7
		IGSM	5
		GTEM REF23	4
		KEI-Linkages	4

*All the versions of MESSAGE in AR5 include MACRO (they are sometimes referred to as MESSAGE-MACRO for that reason).

One of the reasons why the model types are important is because several authors have argued that general equilibrium models, including optimal growth and CGE models, exclude negative costs by construction. Already in 1997, DeCanio argued that the sign of the aggregate costs produced by general equilibrium models is predetermined by the assumption that there *must* be a trade-off between environmental protection and economic growth (DeCanio, 1997). Terry Barker, who played an instrumental role in the development of the macroeconometric E3MG model (which, as shown in section 4.2.2, have predicted net negative costs of mitigation), has long criticised CGE models for assuming that the economy is already in an optimal equilibrium, which implies that climate policies can only have a negative impact on economic output. In CGE models, according to Barker, "by definition any change brought about by policy towards sustainability will incur economic costs" (2004, p. 8). Similar arguments have also been put forth by others (see for example Scrieciu et al. (2013) and Wolf et al. (2016)).

Given the influence of MERGE on optimal growth IAMs (see Chapter 2), and the influence of optimal growth models on AR5 cost estimates, it is worth looking at how MERGE first conceptualised the climate change problem. The part of MERGE that is used to estimate the cost of mitigation, Global 2200, uses production functions to calculate economic output based on capital, labour, and energy. Savings decisions are made so as to maximise the discounted utility of consumption corresponding with the standard Ramsey optimal growth framework. According to Manne et al.,

> "[e]nergy-economy interactions occur in two ways…energy is an input to the economy…energy costs represent one of the claims upon the economy's output. Tighter environmental standards and/or an increase in energy costs will reduce the net amount of output available for meeting current consumption and investment demands." (Manne et al., 1990, p. 57)

Thus, environmental policy by definition reduces the amount of economic output in MERGE: emissions constraints necessarily lower GDP. In other words, the original framework underpinning the MERGE model excludes the possibility of net negative costs by construction.

Although not part of the AR5 IAMs, Nordhaus' contribution to climate economics is also worth considering due to its significant influence on IAMs in general and on optimal growth models in particular. Nordhaus also had a strong influence on the work of Manne, including the MERGE model (Nordhaus, 2017). In his seminal article, "To Slow or Not to Slow", Nordhaus (1991a) writes,

> "[w]e can derive from economic theory certain properties about the shape of the marginal abatement cost function in a competitive economy with no other externalities and where controls are efficiently designed. First, we know that it has a minimum of zero at the uncontrolled point: The first units of GHG reduction are virtually free. This is the result of the zero market price on the GHG emissions. Second, we know that the cost function increases in the level of abatement. Third,

> society can always do worse than the abatement cost function by inefficiently designing regulations". (Nordhaus, 1991a, p. 923)

Thus, according to Nordhaus, the cost of mitigation must also be positive, and it must increase with the level of mitigation. This is indeed consistent with the AR5 cost estimates if we look at the results (Figure 4.1). As stated in the quote, however, these properties assume that two conditions hold: first, that the economy is perfectly competitive and that no other externalities (beyond the climate externality) are present, and second, that climate policies ("controls") are perfectly designed. These two conditions are crucial. In particular, and as already indicated in Table 4.1, the possibility of net negative costs is intimately related to the presence of already existing (non-climate) market failures. In other words, for mitigation costs to be negative, the first condition must fail. Nordhaus (1991a) does not discuss this possibility. He does, however, note that society can "do worse" by inefficiently designing policies, i.e. by failing to meet the second condition. As we will see in section 4.4.5, relaxations of Nordhaus' second condition also play a much larger role in the AR5 IAM ensemble.

The above might suggest that the reason why AR5 does not contain any net negative cost results is because the types of IAMs that are used to produce the cost results exclude such results by construction. Not only have several authors argued that general equilibrium models simply *assume* net positive costs, but the most influential optimal growth model in the AR5 ensemble also appears to have assumed this. Given how later optimal growth models in the AR5 ensemble relate to MERGE (see Chapter 2), we might be tempted to conclude that this is the reason why the AR5 ensemble does not capture the possibility that mitigation costs might be net negative.

However, section 4.2.2 gave examples of both optimal growth (FEEM-RICE-FAST) and CGE (IMACLIM and GEM-E3) models that have produced net negative cost results. Thus, it simply cannot be true that all general equilibrium models exclude the possibility of net negative costs by construction. Claims that they do are incorrect; The real picture is more nuanced.

In order to explain how general equilibrium models can produce net negative cost results, the next section shows how an optimal growth model can be modified to take into account mechanisms that can lead to net negative costs. These modifications are then related to the concrete modelling studies presented in section 4.2.2.

4.4.2 "Optimal green growth" and imperfect baselines

The issue with standard optimal growth models is, according to Hallegatte et al. (2012), that they assume a first-best world with no market failures in which output is a simple function of human capital (A), physical capital (K), and labour (L), $Y = f(A, K, L)$. In this world, all production factors are used optimally. In order to take the environment into account, Hallegatte et al. modify the standard optimal growth framework by including the environment as 'natural capital' in the production function[43], such that $Y = f(A, K, L, E)$, where E is natural capital. The inclusion of natural capital allows the possibility that production factors are not used in an optimal manner when natural capital is under-priced.

The above production function does not, however, take into account other existing market failures and the impacts of environmental policies on the economy (other than via E). In order to represent these factors, Hallegatte et al. (2012) add an efficiency factor (Ψ) (which takes on values between 0 and 1) and environmental policy (P_E) to the production function, such that

$$Y = \psi(P_E) f\big(A(P_E), K(P_E), L(P_E), E(P_E)\big).$$

In this production function, every factor is a function of the environmental policy. These modifications mean that many of the green growth mechanisms listed in Table 4.1 can be captured by the model. For instance:

i. The correction of existing market failures (such as the energy efficiency gap) can be represented as an increase in the overall efficiency factor, ψ. This factor can also capture demand-led and Keynesian effects (which, according to Hallegatte et al. (2012), are only short-term) in situations of high unemployment and underutilized production capital.
ii. The notion that economic output can be increased by environmental policy if we take into account the impacts on aspects such as worker's health and the risk of natural disasters (which affect productivity) can be represented by increasing the effective quantity of K and L.
iii. The possibility of positive knowledge spillovers from accelerated innovation can be represented as an increase in A. This amounts to moving the production frontier.

Thus, an optimal growth framework can be modified to take into account many of the sources of net negative costs, which are reflected in the literature and listed in Table 4.1. Despite the number of mechanisms discussed in section 4.2, however, the possibility that climate policy can have a positive impact on the economy rests on two crucial assumptions: First, the assumption that the economy is not in an optimal state to begin with, and second, the assumption that climate policy can bring the economy closer to its optimal state. If either of these assumptions fail, climate policy can only lead to net positive (or in the best case zero) costs.

[43] This is different from including the environment into the utility function in order to express its so-called amenity value: arguing that the environment is important because it contributes directly to utility is different from arguing that it is a factor of production that contributes to economic output (which again contributes to utility).

The same is true for CGE models (which are responsible for 21% of the scenarios generated by general equilibrium IAMs in AR5). While optimal growth models allocate consumption and investment so as to maximise the discounted utility of a representative household's utility over its lifetime, CGE models adds to this an optimization also of the allocation of resources between different sectors of the economy (Wolf, Schutze, et al., 2016). In either case, if the baseline is optimal, climate policy will come at an economic cost. Essentially, what the modifications applied by Hallegatte et al. (2012) to the above optimal growth framework do, is allow for climate policy to correct for market failures that are present in the baseline.

With this in mind, we can look back at the examples of optimal growth and CGE models that have produced net negative costs presented in section 4.2.2 and get a better understanding of why they were able to do so.

First, FEEM-RICE-FAST assumed market imperfections due to externalities in R&D investment. When climate policy was introduced, regions increased their R&D investments, thereby bringing these investments closer to the optimal level. Thus, an existing market failure was corrected in this model by introducing climate policy. In addition to this, FEEM-RICE-FAST also included learning-by-doing. Second, the version of GEM-E3 that produced net negative costs in Jaeger et al. (2011) allowed labour to be under-employed in the baseline. This meant that climate policy could increase employment, which acted as a stimulus on the economy. This version of GEM-E3 also included learning-by-doing. Lastly, IMACLIM also included learning by doing. In order for IMACLIM to predict net negative costs, however, two additional assumptions were imposed by Jaeger et al. (2015): revenue recycling and the implementation of a stimulus package targeted at low-carbon technology.

In both of these CGE models, it should be noted, decision-making is myopic. That is, agents do not have perfect foresight. This means that these models *in general* don't depict optimal baselines. However, unless these myopic (non-optimal) behaviours are corrected for by the climate policies that are implemented, the lack of foresight in these models often means that costs become *even higher*, not lower. In fact, myopic IAMs are known to predict high costs compared to perfect foresight IAMs (Babiker et al., 2009). Myopic behaviour, however, is also key to negative costs: It opens the possibility that climate policy can correct non-optimal behaviour and thus bring the economy closer to its optimum.

In summary, both optimal growth and CGE models can produce net negative cost results if they are modified such that the baseline is not optimal and if climate policy can correct economic imperfections in the baseline. Several of the green growth mechanisms listed in Table 4.1 play a key role in enabling net negative costs in the modelling examples reviewed in this chapter.

4.4.3 A closer look at the IAMs that produce economy-wide cost estimates in AR5

It follows from the above that not including existing market imperfections in optimal growth and CGE models means that they cannot depict net negative cost results. According to Hallegatte et al., "since the potential for accelerated income growth thanks to green growth policies arises from market failures, they cannot be assessed using models unable to represent these market failures" (Hallegatte et al., 2012, p. 4). Given the lack of net negative cost results in AR5 (Figure 4.1), one might hypothesise that the reason behind this is that most IAMs in AR5 do not represent such market failures.

In order to find out whether this is the case, this section reviews AR5 scenario publications and model documentation to find out whether market imperfections and other negative cost mechanisms are included in the IAMs used to generate the AR5 cost estimates. The imperfections that are considered are informed by the mechanisms identified and discussed in section 4.2 (listed in Table 4.1). The AR5 scenario publications consist of 16 model intercomparison overview publications and two studies that present additional scenarios used to estimate costs in AR5 (Prinn et al., 2011; Riahi et al., 2011). All scenarios responsible for economy-wide cost estimates in AR5 are captured by these publications. The model documentation is the same as the documentation that was gathered in Chapter 2 and used to create the model family tree.

In order to find out whether market imperfections and associated green growth mechanisms are present in the IAMs used to estimate the cost of mitigation in AR5, a search of terms that are typically used to describe such market imperfections and green growth mechanisms was conducted for the scenario publications on the one hand and the model documentation on the other hand. In addition to this, general terms relating to the possibility of net negative costs (e.g. "negative costs" and "economic gain") were searched for. The results are presented in Table 4.3 and Table 4.4. The publications that were included in the searches are listed in Appendix C (which also shows what general equilibrium IAMs were part of what model inter-comparison exercises and what general equilibrium IAMs were not part of any model inter-comparison exercises). The tables indicate only hits where the terms are used in association with market imperfections or green growth mechanisms. That is, the tables do not show hits where the terms listed are used in relation to something else. For example, the term "failure" is frequently used to refer to the failure of the Copenhagen climate summit and the term "net negative" is often used when discussing emissions, but only uses of the terms associated with market imperfections and green growth mechanisms are shown in the table.

Table 4.3 Search of terms associated with net negative mitigation costs and green growth mechanisms in AR5 scenario publications (including 16 AR5 Model Intercomparison overview publications and two other scenario publications).

Terms related to net negative costs	Model hits	In relation to
"win-win"°, "negative cost", "net negative", "GDP gain", "economic gain", "costless", "revenue recycling", "Keynes", "Schumpeter"	No hits*	
"double-dividend"°, "distortion", "no-regret", "sub-optimal"°, "economic benefit", "benefit", "imperfect", "learning", "external", "failure"	**IMACLIM**	IMACLIM includes all of the mechanisms listed (G. J. Blanford et al., 2014; Kriegler et al., 2014; Luderer et al., 2012)
"learning", "external",	**WITCH**	WITCH includes learning-by-doing and accounts for technology externalities generated by R&D (Luderer et al., 2012, 2016)
"learning"	**REMIND**	REMIND includes learning-by-doing (Luderer et al., 2012, 2016)
"learning"	**MERGE-ETL**	MERGE-ETL includes learning-by-doing (Kriegler et al., 2015a)

* Only hits related to the ability of models to depict net negative costs are included. E.g. discussions of "net negative" emissions technologies or" co-benefits" are not considered.

° All hyphenated terms are also searched for without the hyphen.

Table 4.3 shows two things. First, none of the terms associated with net negative cost results in general (such as "win-win", "negative cost", "net negative", "GDP gain", and "economic gain") are found in any of the AR5 scenario publications. This shows that the possibility of net negative costs (or economic gains) was not discussed explicitly in the scenario publications, and it indicates that it was not considered in the model intercomparison studies that generated most of the scenarios in the AR5 ensemble, a finding that will be further corroborated in section 4.4.5. This finding is confirmed by carefully going through the presentations of cost results in the AR5 scenario publications. Figures and tables presenting economy-wide cost estimates can be found in eight of the 19 publications. In none of these is the possibility of net negative costs or the absence of such results mentioned or discussed. Most of the intercomparison overview sources do, however, discuss increases in mitigation costs due to either delayed or fragmented action, or limited technology availability (section 4.4.5 will discuss this in more detail).

Second, for the terms that are used to capture net negative cost mechanisms, almost all appear in relation to one AR5 IAM only, namely IMACLIM. For IMACLIM, several publications discuss a number of features that are associated with green growth mechanisms, including existing distortions, sub-optimal baselines, no regret potentials, and the possibility of double dividends and economic benefits. The fact that IMACLIM features imperfect foresight is also frequently referred to (even though this is the case for several other IAMs in AR5 (see Table 3.1), this is not highlighted in these publications). The emphasis on the inclusion of such features in IMACLIM by multiple authors, makes IMACLIM stand out. Recall that IMACLIM is also one of the CGE models that were used to demonstrate the possibility of net negative costs in Jaeger et al. (2015). The fact that IMACLIM, in AR5 produced only positive cost results is discussed in the next section. The only two mechanisms that are mentioned in relation to other AR5 IAMs is learning-by-doing and innovation externalities. Learning-by-doing is included in REMIND and MERGE-ETL, and WITCH includes both learning-by-doing and innovation externalities (from R&D). The fact that none of the mechanisms (listed in Table 4.2) appear to be included in any of the other IAMs in AR5 indicates that these IAMs do not represent market imperfections and sub-optimalities to a significant degree (if at all). The inclusions only of learning and innovation externalities in WITCH, REMIND, and MERGE-ETL, might also be insufficient for generating net negative cost results, as discussed further below.

In order to find out whether market failures and associated green growth mechanisms might have been included in other versions of the same IAMs, however, the model documentation collected in Chapter 2 – which captures the entire lifespans of the 14 most influential IAMs (eight of which are general equilibrium models) – is also analysed. Since model documentation covers technical aspects of individual IAMs in much more detail, one might also expect information about market failures and green growth mechanisms that may not be listed in scenario publications to be present here. The results are shown in Table 4.4.

Table 4.4 Search of terms associated with net negative cost results and green growth mechanisms in the model documentation for the eight most influential of the general equilibrium IAMs in AR5.

Terms related to net negative costs	Model hits	In relation to
"net negative", "GDP gain", "economic gain", "distortion"	No hits*	
"double-dividend"°, "no-regret"°, "negative cost", "economic benefit", "sub-optimal"°, "revenue recycling", "imperfect", "learning"	**IMACLIM**	IMACLIM allows revenue recycling as an option (Kriegler et al., 2015a). Imperfect foresight (Crassous et al., 2006). Other (Bibas & Méjean, 2014; Kriegler et al., 2015a; Waisman et al., 2012).
"revenue recycling"	**GEM-E3**	The model allows revenue recycling as an option (Kriegler et al., 2015a).
"revenue recycling"	**SGM**	The model allows revenue recycling as an option (Brenkert et al., 2004).
"imperfect"	**Worldscan**	Dedicated versions, which extend the core version, that take into account R&D spillovers and imperfect competition, exist (Kriegler et al., 2015a)
"no-regret"°, "costless"	**MERGE**	Inclusion of "no-regrets" options, but with a limit to their extent (Manne et al., 1995). Exclusion of "free lunch" (Manne & Richels, 1990).
"learning", "external"	**MERGE-ETL**	MERGE-ETL includes learning-by-doing for several technologies and thus assumes positive externalities from knowledge (S. Kypreos & Bahn, 2003; Socrates Kypreos, 2007).
"negative cost", "learning", "external", "failure"	**WITCH**	Exclusion of zero and negative cost options, but inclusion of learning-by-doing and accounting for positive externalities due to R&D, i.e. representation of innovation market failures (Bosetti et al., 2009).

“sub-optimal”°, “imperfect”, “learning”, “learning”	**REMIND**	Learning-by-doing is included in REMIND (Bauer, Baumstark, et al., 2012), but exclusion of sub-optimalities in the baseline in the default setting (Luderer et al., 2011). Imperfections in the baseline are not analysed (Bauer et al., 2012). I.e. the model solves for the optimal solution in the absence of externalities (Leimbach; et al., 2010).

* Only matches related to the ability of models to depict net negative costs are reported. E.g. discussions of “net negative” emissions technologies or non-climate ancillary “benefits” are not counted.

° All hyphenated terms are also searched without the hyphen.

Table 4.4 again shows that generic terms associated with net negative cost results and win-win opportunities (“net negative”, “GDP gain”, and “economic gain”) never appear in the model documentation for the eight most influential of the general equilibrium IAMs in AR5. Table 4.4 also confirms the finding in Table 4.2 for IMACLIM, namely that it includes a number of features associated with the possibility of net negative cost results. In addition to IMACLIM, however, seven other IAMs also appear in Table 4.3. Among other things, it is found that revenue recycling is an option in the two CGE models GEM-E3 and SGM, and that dedicated versions that incorporate multiple market imperfections exist for the CGE model Worldscan. Although, judging by the publications that are specific to AR5 (presented in Table 4.3), these features are not included in the versions of these IAMs that were used to generate the scenarios that are included in the AR5 ensemble, we cannot know for sure whether or not these possibilities are included or excluded from the corresponding AR5 scenarios. When it comes to the four optimal growth models that are listed in Table 4.3 (MERGE, MERGE-WTL, WITCH, REMIND), both the inclusion and the exclusion of green growth possibilities associated with the terms are found. In MERGE, although some level of no-regret options is included, the overall assumption is that “there is no free lunch” (A. S. Manne & Richels, 1990). In the case of WITCH, zero and negative cost options (i.e. no-regret options) are excluded (Valentina Bosetti, Tavoni, De Cian, & Sgobbi, 2009), but the model incorporates market failures associated with R&D and learning-by-doing. As already noted in Table 4.3, MERGE-ETL and REMIND also include learning-by-doing.

In summary, thus, IMACLIM appears to include a range of mechanism that could lead to net negative cost results. The three CGE models, GEM-E3, SGM and WorldScan, are set up in such a way that they allow for the representation of revenue recycling and spillovers, although there are no signs of the inclusion of these options in the AR5 scenario documentation (Table 4.3). When it comes to the three optimal growth models, MERGE-ETL, WITCH and REMIND, these incorporate knowledge externalities caused by learning-by-doing and, in WITCH, also by R&D, but there is no sign of other green growth mechanisms. MERGE, according to early model documentation (which may or may not still be accurate), explicitly excludes the possibility of a "free lunch".

An important question is whether the inclusion of learning-by-doing and R&D is sufficient to allow for net negative cost results in MERGE-ETL, WITCH, and REMIND. Learning-by-doing is primarily seen as a source of inter-temporal market failure: imperfect capture of future payoffs from current actions (Gillingham & Sweeney, 2010). In IMACLIM, because agents don't have perfect foresight in the baseline (or in any other scenarios), climate policies can correct for this imperfection. That is, policies can push agents' behaviour in the direction of what they would have done had they had perfect foresight. Thus, climate policy can have a positive impact on economic output. In MERGE-ETL, REMIND, and WITCH, however, agents have perfect foresight. This means that future cost reductions will already be taken into account in the baselines. Thus, the inter-temporal aspect of the innovation market failure is removed and the incorporation of learning-by-doing does not necessarily lead to a sub-optimal baseline, unless other aspects mean that it does.

In REMIND, it appears, the effects of learning-by-doing are already taken into account in the baseline, meaning that the baseline is already optimal (Luderer et al., 2012). Thus, the inclusion of learning-by-doing in REMIND (although it might lower the cost of mitigation) cannot lead to net negative mitigation costs; knowledge spillovers are already internalized in the baseline and there is no room for improvement. In WITCH, similarly, agents also make inter-temporally optimal decisions in the baseline. However, WITCH also accounts for non-cooperative behaviour between regions. In addition to the inter-temporal aspect of the innovation market failure, learning-by-doing can also lead to underinvestment in new technologies if there is significant knowledge spillover to other actors (Luderer et al., 2012). Regions, in this case, are more likely to free ride on other regions and wait until sufficient global deployment (paid for by others) has taken place for technologies to become cheaper. This effect is indicated by Bosetti et al. (2006) for the WITCH model. That is, WITCH considers positive externalities from learning-by-doing (and R&D) and thus allows climate policies to have a positive impact on economic output. When it comes to MERGE-ETL, all model publications run baselines with and without learning and show that costs and emissions decline when learning is taken into account. Although it is not clear whether the AR5 MERGE-ETL baselines includes learning-by-doing (Kriegler et al., 2015), it would appear inconsistent to compare mitigation scenarios that include learning-by-doing with baselines that exclude learning-by-doing. Thus, in all likelihood, learning-by-doing in MERGE-ETL is not sufficient to allow for net negative mitigation costs.

Overall, the analysis of the AR5 scenario publications and the analysis of the model documentation for the most influential of the general equilibrium IAMs in AR5 suggests that only IMACLIM includes existing market imperfections and associated green growth mechanisms to a significant extent. These findings give reason to believe that the mechanisms that enable net negative cost results in optimal growth and CGE models are either not present, or only present to a limited extent, in most of the IAMs used to estimate economy-wide costs in AR5.

4.4.4 Modifying is difficult, and few general equilibrium models do

The previous section suggests that the general equilibrium IAMs in AR5 include market imperfections and sub-optimalities only to a limited extent. One of the reasons for this might be that there are no general rules for making such modifications. According to Krugman

> "we have a body of economic theory built around the assumptions of perfectly rational behavior and perfectly functioning markets. Any economist with a grain of sense — which is to say, maybe half the profession? — knows that this is very much an abstraction, to be modified whenever the evidence suggests that it's going wrong. But nobody has come up with general rules for making such modifications." (Krugman, 2014).

In short, using general equilibrium models to depict the possibility of net negative costs appears to require what we might call "non-standard" assumptions. Introducing non-standard assumptions requires justification. In applied IAMs it also requires data. Krugman writes that what economists "really do is combine maximization-and-equilibrium as a first cut with a variety of ad hoc modifications reflecting what seem to be empirical regularities about how both individual behavior and markets depart from this idealized case" (Krugman, 2014). One of the problems for climate economics, however, is that there are few empirical regularities to draw from. It is not possible to empirically test whether we can have net economic benefits from climate mitigation at the level necessary to meet the Paris target because data on the economic consequences of mitigation at this level is not yet available (Scrieciu et al., 2013)[44]. Although the existence of a number of market failures, such as positive knowledge spillovers from technology R&D and under-investment in energy efficiency improvements, is apparent, and even though the existence of these market failures clearly indicates that the current situation is not optimal, we have no empirical data on the role of such market failures during a complete transformation of the energy system.

Without agreed upon rules and with only limited empirical evidence, the inclusion of non-standard assumptions in models can easily be perceived as arbitrary and unfounded. Under such circumstances,

[44] Rosen and Guenther (2015) also point out that since we can never know what the cost of the counterfactual baseline would have been, we will never be able to determine the net economic benefits of mitigating climate change, even in hindsight.

it might be safer to stick to standard assumptions, which in the case of optimal growth and CGE models imply optimal baselines and perfectly competitive markets that maximize consumer and producer surplus. If modelers lack both data and an accepted framework in which to incorporate existing sub-optimalities into their models, they might conclude, even though such sub-optimalities are known to exist, that it is not epistemically justifiable to do so.

In fact, if we look back at the examples of model studies that predicted net negative costs, we find clear indications that the necessary modifications were difficult to make. According to the study, *A New Growth Path for Europe*, the feedbacks between investment, learning-by-doing, and expectations that allow net negative cost results "are hard to implement in existing models and have been neglected so far" (Jaeger et al., 2011, p. 23). This study was based on a concerted effort to *enhance* the GEM-E3 model in order to allow for the identification of potential win-win strategies. According to Jaeger et al., "developing enriched models along these lines is *a major research program that will keep many researchers busy for many years*" (2011, p. 25 my italics). According to IMACLIM modelers, also, "it remains hard to identify and assess market imperfections" (Guivarch et al., 2011, p. 2).

It would make sense that it is easier to get hold of evidence and empirical data that can support non-standard assumptions when studies are limited to a specific time and place. According to Hallegatte et al., "one can identify channels that are theoretically able to make green policies contribute to economic growth; however, detailed and country- and context-specific analyses of each of these channels are required to conclude that this will in fact happen in any specific situation" (2012, p. 4). It is telling that both of the Jaeger studies focused on Europe only, and that both of these studies were also conducted in the aftermath of the great financial crisis, which reduced European GDP by several percentage points (Jaeger et al., 2011). Among other things, Jaeger et al. write, "in the current European situation, there is evidence for a suboptimal use of capital" (2015, p. 56). Thus, it seems, the sub-optimal nature of the situation was particularly obvious in these studies.

Even though the two studies by Jaeger et al. demonstrate that it is possible to generate net negative cost results using CGE models (notably GEM-E3 and IMACLIM, both of which are also included in the AR5 ensemble), these studies still appear to be unique. In particular, these studies were conducted in a context in which European climate policy models (which are primarily CGE models) were seen to be in need of major adjustments. According to Jaeger et al. "the experience of the global financial crisis shows that the existing economic models were seriously limited. Against this background, a fundamental overhaul of European climate policy models is required" (Jaeger et al., 2011, p. 6). In fact, Jaeger et al. write, "the canonical models of climate economics [which includes optimal growth and CGE models]…exclude win-win options by design. For climate policy, however, it is exactly these options that matter" (2011, p. 39). Thus, Jaeger et al. end up mirroring the claims made by authors such as DeCanio (1997) and Barker (2004). Clearly, GEM-E3 and IMACLIM are, according to Jaeger et al. exceptions to the rule: "for the first time in the academic climate modeling field, the present study has taken a state-of-the-art model of climate economics and enhanced it…" (2011, p. 6). Two years after this study was published (and two years before the AR5 WGII report was published), Scrieciu et al. (2013, p. 256) still note that "CGE models that allow for externalities and 'inferior' equilibria in their depiction of economies are extremely sparse in the literature."

When it comes to IMACLIM, the previous section already indicated that this model stands out when it comes to the inclusion of mechanisms that can enable net negative cost results. Recall that, while GEM-E3 was specifically enhanced for the Jaeger et al. (2015) study, IMACLIM was not (although particular policy assumptions had to be imposed in order for IMACLIM to produce net negative rather than net positive cost results). According to IMACLIM modelers themselves, "almost all numerical models used in climate policy studies assume a perfect labour market and neglect unemployment issues, even those complex computable general equilibrium models whose comparative advantage is intended to represent subtle macro feedbacks" (Guivarch et al., 2011, p. 4). The IMACLIM cost profiles, which depict short-term losses followed by long-term catch-up, are, according to Waisman et al., also "non conventional" (2012, p. 109) and characterised by a "peculiar shape…compared to those in most models, including WITCH and REMIND" (2012, p. 102).

When it comes to the reasons why IMACLIM does not produce net negative cost results in AR5, two things should be noted. First, as already noted, specific policy assumptions were imposed in the Jaeger et al. (2015) study in order for IMACLIM to predict net negative as opposed to net positive costs. These assumptions included revenue recycling and a targeted stimulus package. In AR5, only carbon taxes are used to reduce emissions. This might be one of the reasons why IMACLIM does not predict net negative costs in AR5 (but in the Jaeger et al. (2015) study). Second, as noted by Edenhofer et al., IMACLIM "adopts a pessimistic view of technological change by assuming strong inertia and by neglecting carbon-free energy sources from backstop technologies" (2006, p. 68). Among other things, IMACLIM "show[s] that taking into account labour market imperfections leads to higher macroeconomic costs of climate mitigation than in the case of perfect labour markets" (Guivarch et al., 2011, p. 12). This leads us to a crucial point: assuming a non-optimal baseline does not automatically imply that climate mitigation represents a win-win option. Whether it does or not depends on a plethora of assumptions and policy options. Under certain assumptions of the labour market (very low absolute values of the wage curve elasticity), for example, and using a relatively low discount rate[45] (3%), the

[45] The discount rate is particularly important for the NVP of mitigation costs in IMACLIM because IMACLIM shows high short-term costs of mitigation but low or even negative long-term costs.

mitigation cost in IMACLIM can also become net negative (Guivarch et al., 2011).[46] Assumptions that baselines *are* optimal, however, nonetheless excludes the possibility of win-win opportunities. In short, as Waisman et al. write, “one can argue that imperfect foresight, incomplete markets and institutional failures will lead to higher costs that those reported so far, or, conversely, that non optimal baselines offer opportunities for relative gains under climate policy” (Waisman et al., 2012, p. 102). Both are possible.

In summary, the CGE models that have been shown to predict net negative cost results, IMACLIM and GEM-E3, appear to represent exceptions from the rule. GEM-E3 had to be specifically enhanced in order to be able to depict win-win possibilities and IMACLIM is by all accounts an outlier among CGE models. Before discussing the roe of such examples, the next section shows that, contrary to enabling IAMs to depict net negative costs, the model intercomparison studies that generated the vast majority of the scenarios in the AR5 ensemble focused on aspects that can only increase the cost of mitigation.

4.4.5 AR5 model intercomparison studies focus on increasing costs

While the estimates of the cost of mitigation are presented in the AR5 Synthesis Report and the WGIII SPM, the most detailed discussion of these estimates is found in section 6.3.6.5 of the AR5 WGIII report. This section notes that the reported costs (shown in figure 4.1) “have assumed an *idealized policy implementation* and in many cases an *idealized implementation environment* with perfectly functioning economic markets devoid of market failures, institutional constraints, and pre-existing tax distortions” (IPCC, 2014a, p. 455 my italics). These two assumptions map almost perfectly onto the two conditions highlighted by Nordhaus in 1991, quoted in section 4.4.1: a competitive economy with no other externalities, and efficiently designed policies. ‘Idealized policy implementation’ describes the situation in which policies are efficiently designed and an ‘idealized implementation environment’ means that there are no other existing market failures. When policies are not implemented efficiently, the idealized policy implementation assumption fails. In this case, costs can only be higher (by definition). If the implementation environment is not idealized, this means that there are other existing market failures. In this case, as indicated above, the impacts on costs can go in both directions. As we have already seen in this chapter, however, capturing the existence of other market failures is crucial for arriving at net

[46] Waisman et al. (2012) also show that the adoption of certain policies in the transport sector leads to a significant reduction in policy costs and even negative costs in the long term. Waisman et al., however, refer to annual costs, i.e. the difference in GDP between the baseline and the mitigation scenario in every year. Saying that annual costs in a mitigation scenario might be net negative is different from saying that the total (discounted) costs in a mitigation scenarios might be net negative (the latter might be positive or negative when annual costs in some years are positive and in other years are negative).

negative cost results. Thus, assuming an idealized implementation environment excludes the possibility of net negative costs.

In order to increase realism, the AR5 scenario ensemble includes a large number of ‘non-idealized policy implementation’ scenarios. The outputs of these scenarios are reported in separate figures and tables in the WGIII SPM and the Synthesis Report just after the idealized scenario results (shown in figure 4.1). If we look closely, we find that 86% of the scenarios in the AR5 ensemble stem from model intercomparison studies that are focused explicitly on non-idealized implementation in the form of delayed or fragmented climate action (or both)[47]. Delayed action implies that “mitigation is not undertaken ‘when’ it would be least expensive” and fragmented action implies that “mitigation is not undertaken ‘where’ it is least expensive” (IPCC 2014 p. 421). Both departures from idealized assumptions necessarily lead to an increase in mitigation costs compared to the idealized case. In short, they represent ways in which “society does worse” by not implementing climate policies efficiently, as noted by Nordhaus. In addition to these departures, many of the model intercomparison studies in AR5 also explore the effects of limited technology availability[48]. Because limiting technology availability necessarily implies a reduction in mitigation options, this also leads to an increase in the cost of mitigation by necessity. In total, this means that eight out of the nine model intercomparison studies, responsible for 88% of the scenarios in AR5, focus on either ‘non-idealized policy implementation’ or limited technology availability[49] or both. All necessarily lead to an increase in the cost of mitigation.

While the potential effects of a ‘non-idealized implementation environment’ is recognised in AR5, it seems few scenarios include such departures. Certainly, none of the model intercomparison studies focus on this.
The potential impacts of assuming a ‘non-idealized implementation environment’ are only discussed briefly in section 6.3.6.5 of the WGIII report:

---

[47] Model intercomparison studies investigating the effects of delayed and/or fragmented action include EMF 22 (Clarke et al., 2009), EMF 27 (Kriegler et al., 2014), AMPERE (Kriegler et al., 2015a; Riahi et al., 2015a), LIMITS (Kriegler et al., 2013), RoSE (Chen et al., 2016), POeM (Lucas et al., 2013), and RECIPE (Luderer et al., 2012).
[48] Model intercomparison studies that investigate the effects of limited technology (sometimes called ‘technology failure’) include EMF 27 (Kriegler et al., 2014), ADAM (Edenhofer, Knopf, Leimbach, & Bauer, 2010), RECIPE (Luderer et al., 2012), and AMPERE (Kriegler et al., 2015a; Riahi et al., 2015a). These multi-studies are together responsible for 65% of the scenarios in the AR5 ensemble.
[49] The only model intercomparison study that did not focus on these issues, AME (Calvin et al., 2012), explores the role of Asia in mitigation.

> "Climate policies will interact with pre-existing policy structures as well as with other market failures beyond the market failure posed by climate change — that is, a non-idealized implementation environment — and these interactions can either increase or decrease policy costs. A number of authors have argued that costs could be much lower or even negative compared to those produced by studies assuming idealized policy and implementation environments (Bosquet, 2000; Bye et al., 2002; Waisman et al., 2012). The results of these studies rest on one or several assumptions — that mitigation policy be used not only to address the climate externality, but also to achieve other policy priorities such as sustainable development; the use of mitigation policy instruments for the correction of the implementation environment including removal of market failures and pre-existing distortions; and/or on optimistic views of climate-related innovation and technology development, adoption, and penetration." (p. 456).

Despite noting the potential impact of such assumptions, however, section 6.3.6.5 states that "many models represent some of these distortions, but most models represent only a small portion of possible distortions and market failures" (IPCC 2014, p. 455). Given the importance of pre-existing market failures for net negative cost results, this statement gives further reason to believe that most of the IAMs in AR5 are unable to depict situations in which climate policies imply economic gains. In addition to this, none of the model intercomparison studies that generated scenarios for AR5 focused explicitly on representing aspects of 'non-idealized implementation environments'.

Ultimately, section 6.3.6.5 acknowledges, "the reality that assumptions of idealized implementation and idealized implementation environment will not be met in practice means that real-world aggregate mitigation costs could be very different from those reported here" (IPCC, 2014a, p. 455). Given the argument put forth in this chapter regarding the importance of the uncertainty of the cost of mitigation, this is a crucial point. Yet, this caveat is only stated once deep inside Chapter 6 of the WGIII report, not in the SPM nor in the AR5 Synthesis Report, where mitigation costs are presented to the wider public[50]. Thus, this caveat will most likely be missed by the majority of those who use the cost of mitigation reported in AR5. Furthermore, even though section 6.3.6.5 acknowledges that costs in principle could go both ways, AR5 overall gives the opposite impression. The scenarios that assume 'idealized policy implementation' are "referred to as 'cost-effective' scenarios, because they lead to the *lowest aggregate global mitigation costs* under idealized assumptions about the functioning of markets and economies" (IPCC, 2014a, p. 421 my italics). Although a trained eye might realise that these scenarios only

[50] The closest we get to such an acknowledgment in the SPM is a footnote stating that "projections from all models can differ considerably from the reality that unfolds" (IPCC, 2014a, p. 10).

represent the lowest cost *given* idealized assumptions, which implies that real costs can go both ways, it clearly suggests that costs can only be higher. The language used by AR5 to describe IAM scenarios strongly indicates that mitigation costs in reality will be higher than those reported, not lower. This is especially true in the SPM, where the AR5 scenarios are "used as a cost-effective benchmark for estimating macroeconomic mitigation costs" (IPCC, 2014a, p. 15). The technical summary adds to this that "substantially higher cost estimates have been obtained based on assumptions about less idealized policy implementations and limits on technology availability" (IPCC, 2014a, p. 57). Overall, this sends a clear message in AR5 that we should not expect anything but positive, and likely higher than estimated, mitigation costs.

4.4.6 Bias in AR5

This section has shown that all the economy-wide cost estimates in AR5 are based on optimal growth or CGE IAMs. Although it is incorrect to say that these models exclude net negative costs by construction, these models cannot produce net negative costs if the baselines are optimal, which appears to be the case when "standard assumptions" are adopted. To depict non-optimal baselines, market imperfections and other sub-optimalities have to be included in optimal growth and CGE models. From an analysis of the AR5 scenario publications and the model documentation for the most influential IAMs, this section has shown that there are few signs that such modifications have been made in the IAMs in AR5 to a significant degree. The main exception to this is IMACLIM, which, under certain conditions have been shown to predict net negative costs. In addition to this, WITCH allows climate policies to correct innovation externalities. The exact reasons why these two IAMs in AR5 generate only net positive cost results requires further investigation.

Overall, very few optimal growth and CGE models in AR5 appear to be modified in ways that enable net negative cost results. Although it is incorrect to say that all general equilibrium IAMs exclude net negative costs by construction, this section gives reasons to believe that many general equilibrium IAMs exclude net negative costs *in practice*. Basing cost estimates exclusively on general equilibrium models, thus, appears to introduce a bias against net negative cost results. This might be explained in part by the fact that such modifications are difficult to implement. The examples presented in this chapter of optimal growth (FEEM-RICE-FAST) and CGE (IMACLIM and GEM-E3) models that have predicted net negative costs appear to represent the exceptions rather than the rule.

This implies that, although model frameworks on their own do not determine cost results – every IAM output depends on a range of assumptions, many of which can be varied within and across frameworks – model frameworks impact the direction of cost results. A broader diversity of modelling approaches in IAM ensembles will, everything else being the same, reduce the risk of being wrong.

Lastly, this section has also shown that the AR5 IAM ensemble depends heavily on the model intercomparison studies that were conducted after AR4. This means that the focus of these model intercomparison studies – and the assumptions that are explored in these – have a significant impact on the cost estimates in AR5. While AR4 model intercomparison studies focused on the role of induced technological change (Edenhofer et al., 2006), which is widely known to reduce the cost of mitigation (Grubb et al., 2006), AR5 model intercomparison studies focused on the effects of delayed and fragmented action and limited technology availability, aspects that are known to increase the cost of mitigation. Thus, not only model frameworks, but also model intercomparison exercises represent potentially important sources of ensemble bias.

4.5 Conclusion

This chapter has argued, based on a recently revitalised debate in the philosophy of science, that the possibility that the cost of mitigation might be negative should be taken seriously. Even if we believe that net negative mitigation costs are unlikely, their possibility should be taken into account in IPCC reports because being wrong about the sign of the mitigation cost could have large negative consequences. This comes in addition to the mandate that the IPCC already has to communicate the strength of and uncertainties of findings. In some ways, it explains exactly why this mandate is important: given the high stakes involved in climate policy, ignoring the uncertainties of findings implies a significant risk of being wrong. The literature reviewed in this chapter suggests a possibility that mitigation costs might be net negative; experts disagree. This, however, is not reflected in the AR5 results that are provided in the Synthesis Report and the WGIII SPM.

All the cost estimates in AR5 are generated by optimal growth and CGE IAMs, both of which go under the umbrella of general equilibrium models (as seen in Chapter 3). Although it would be incorrect to say that general equilibrium models cannot generate net negative costs by construction, this chapter suggests that the modifications that are needed to enable net negative cost results in these models are difficult to implement in practice. A review of the AR5 scenario publications and the model documentation for the most influential IAMs in AR5 indicates that such modifications are only implemented to a significant degree in one CGE IAM (IMACLIM) and to some degree in one optimal growth IAM (WITCH). This finding confirms what is stated inside Chapter 6 of the AR5 WGIII report, namely that most of the IAMs used to assess the costs of mitigation "represent only a small portion of possible distortions and market failures" (IPCC 2014, p. 455).

Given that the general equilibrium models that have been used to produce net negative cost results reviewed in this chapter appear to represent either "non-conventional" models (IMACLIM), to impose "optimistic" assumptions (FEEM-RICE-FAST), or to be specifically "enhanced" for the purpose of identifying win-win opportunities (GEM-E3), one might wonder if the net negative cost results are realistic. Do the assumptions and policies imposed by Jaeger et al. (2011, 2015) not amount to cherry picking? In these studies, GEM-E3 was specifically enhanced and IMACLIM was run (using the default setup) under very specific policy assumptions to depict pathways in which emissions reductions produce net economic gains. Two things can be said in response to this: First, as long as the assumptions are plausible, the results indicate real possibilities[51]. Second, the fact that some of the policy assumptions are chosen specifically in order to arrive at net negative costs does not pose a problem as long as these assumptions reflect real policy choices. Policies, after all, represent control (independent) variables in policy models: If IAMs show that revenue recycling can lead to net economic benefits of climate mitigation, this is an insight that can be used to inform the design of climate policies in such a way that net negative costs can be realised. The point in this chapter is not that net negative costs are more likely than net positive costs but that net negative cost – according to the literature and applied modelling studies – represent a real possibility that is excluded from AR5 results and that this possibility is *important*. A direct implication of this is that we don't simply want to provide best guess estimates of the cost of mitigation, but that the tails – i.e. unlikely but nonetheless plausible outcomes – are also important. By extension, non-conventional general equilibrium IAMs – as well as "optimistic" and "pessimistic" assumptions (as long as these are plausible) – serve an important purpose and deserve a place in IPCC reports.

Although optimal growth and CGE IAMs can depict win-win possibilities given certain assumptions, the ability to use such IAMs to depict win-win possibilities hinges on our ability to first identify, and secondly implement in the models, any market imperfections that may be removed with climate policies. For macroeconometric IAMs, the situation is different, because optimality does not enter in the same way (the difference between the baselines and the mitigation scenarios depend on the econometrically estimates equations, which are based on historical data). For the same reason as above, IAMs (outside the general equilibrium category) that are seen as non-conventional should – as long as these are based on plausible and theoretically justified assumptions[52] – also be included in IPCC reports, to the extent possible, in order to capture the full uncertainty of the cost of mitigation. Two of the models

[51] Unless, of course, one doubts the models themselves, in which case it doesn't matter what the assumptions are.

[52] There will of course always be debate and disagreement within the economics and integrated assessment modelling community regarding what is plausible and implausible, and what is and isn't theoretically justifiable. But, shifting the burden of proof from those who want to *include* certain IAMs to those who want to *exclude* certain IAMs might help increase the diversity of IAM ensembles, which was argued in Chapter 2 to be important for our ability to draw robust insights, and which is argued in this chapter to be important in order to capture the uncertainty of key IAM outputs.

that have predicted net negative costs reviewed in this chapter are not general equilibrium models. T21 is a systems dynamics model and E3MG is a macroeconometric model. No model types of this sort are included in the AR5 ensemble[53]. Whereas costs in AR5 are exclusively positive, the inclusion of E3MG in AR4 contributed to a picture in which the cost of mitigation could be both positive and negative. While it is beyond the scope of this chapter (and thesis) to examine the processes by which IAMs are included or excluded from IPCC reports, the argument presented in this chapter suggests that this is an important area for future research. As of now, IIASA and the Integrated Assessment Modelling Consortium (IAMC) are responsible for putting together the IPCC scenario databases and checking that scenarios meet criteria. How those criteria are chosen and what they imply for the IAMs that are included in IPCC reports is, however, not clear.

In terms of what the IPCC can and should do, the options are similar to those laid out in the conclusion to Chapter 2. First, the IPCC should acknowledge and discuss the uncertainty of mitigation cost estimates in synthesis reports and the SPMs. A few sentences within one chapter in the WGIII report is not sufficient for this message to be taken up by the wider public and policymakers. Second, the goal of including a variety of different perspectives could be built into the process by which IAMs are selected for inclusion in IPCC reports. Identifying IAMs that stand out is not difficult. The publications and documentation for not only IMACLIM, but also E3MG and T21, explicitly present these models as different from most other macroeconomic models (Barker & Şerban Scrieciu, 2010; UNEP, 2011). Third, the IPCC could use their authority to call for more research into e.g. the feasibility and the likelihood of net negative mitigation costs. This could be used to provide a better picture of the tails of the cost distribution. Additionally, this might also incentivise modelers to expand the approaches commonly taken and include non-standard assumptions (and frameworks). Fourth, this chapter has shown that one of the issues with AR5 is its dependence on model intercomparison studies that focus on aspects that can only increase the cost of mitigation. A fourth option, if the goal is to reflect the "true" uncertainty of cost estimates, could therefore be for IPCC reports to not only include results that are produced in between assessment reports, but to also include the results from previous model intercomparison studies. In that way, IPCC cost estimates could be based on a larger set of feasible assumptions, which is also a precondition for robust insights.

Finally, IAM modellers can better communicate the assumptions that determine cost results, the ability of their models to produce net negative (and net positive) values, and their level of confidence and reasons for making said assumptions. According to the argument in this chapter, integrated assessment

[53] There is one IAM in the AR5 ensemble that is described as 'econometric' (see Table 3.1). This IAM, however, does not contribute to the economy-wide cost estimates in AR5.

modelers have a moral responsibility for the consequences of making inaccurate or unreliable claims. Better communicating the conditions under which their claims are thought to hold, would be a good place to start to increase their accuracy.

5 A different simulation model: introducing FTT

Reaching the Paris target requires a rapid turnaround in emissions trends. As seen in Chapter 1, CO_2 emissions need to reach net zero around 2050 to limit global warming to below 1.5˚C and 2070 to limit global warming to below 2˚C[54]. A key question is therefore whether emissions can be reduced quickly enough to limit global warming to "well below 2˚C". This question usually contains within it a question of whether there is or will be sufficient political will to implement necessary policies. However, even if we ignore the (crucial) question of political will, there is a question of the impacts of policies on the speed and magnitude of technological change. What is the impact of climate policies on the energy system, and related emissions, and how reliably can we predict it?

So far, this thesis has argued for the importance of a diversity of modelling approaches. This chapter and the next focus on the use of the Future Technology Transformations (FTT) model – a relatively new energy system model that is seen as different from most energy system models – to predict the impact of climate policies on emissions. While Chapter 6, which presents the brunt of the work on FTT, analyses the sensitivity of FTT predictions to parameter values, this chapter provides the context and motivation for so doing. One of the main reasons for choosing the FTT model is that it incorporates features that are seen to distinguish it from more widely used energy system models based on optimisation (ESOMs) such as MESSAGE and MARKAL/TIMES. Importantly, FTT is seen by Mercure et al. to offer a "a more realistic modelling approach" (2016, p. 102) than what ESOMs generally do. This and other features that, according to FTT modelers, represent improvements compared to ESOMs are presented in this chapter. The point of so doing is not to argue that FTT is superior or "better" than ESOMs, nor is it to criticise ESOMs, but to examine whether FTT meets the goals that it (according to those who built it) was designed to meet. Despite neoclassical assumptions and exogenous technology deployment constraints posing issues for ESOMs, this chapter argues, we cannot claim that FTT offers more reliable predictions of policy impacts without (at least) assessing the sensitivity of said predictions to assumptions. This is what Chapter 6 begins to do. Although relatively comprehensive sensitivity analyses have been conducted for some other IAMs (e.g. Bosetti et al. (2015), Rogelj et al. (2013), Iyer, et al. (2014), Barron and McJeon (2015), Olaleye and Baker (2015), McJeon et al. (2010), Lemoine and McJeon (2013), Lehtveer and Hedenus (2015)), no comprehensive sensitivity analysis of the kind performed in Chapter 6 has so far been conducted for FTT. As such the analysis in Chapter 6 offers a first step towards and improved understanding of the behaviour of the FTT model and the sensitivity of its predictions to parametric assumptions.

[54] This corresponds to a 50% chance of staying below 1.5˚C and a 66% probability of staying below 2˚C.

Section 5.1 presents the FTT model and section 5.2 discusses the distinction between optimisation and simulation in the climate policy and modelling literature more generally. Section 5.3 explains how ESOMs have been used both to identify optimal policies and to predict the evolution of energy systems and presents some of the assumptions such predictions rely on. Section 5.4 presents the two features that FTT was designed with in order to increase the realism of predictions of policy impacts: a fine-grained representation of policies and an endogenous derivation of technology deployment rates based on diffusion theory. Focusing on the latter, section 5.4 then explains why technology deployment constraints can have a strong influence on scenarios generated by ESOMs, but how related assumptions are no less important for FTT. Section 5.5 concludes.

5.1 Future Technology Transformations (FTT)

FTT was designed by Jean-Francois Mercure at the Cambridge Centre for Climate Change Mitigation Research (4CMR)[55] in the Department of Land Economy in order to predict the impacts of climate policies on technology diffusion. Understanding the impacts of policies on technology diffusion is crucial because it determines the future technology mix, which is a key determinant of emissions. FTT was designed specifically to capture technology dynamics, including learning, at a global level.

FTT currently consists of four different sub-models, capturing the power, passenger vehicles, steel production, and household heating sectors[56] (Cambridge Econometrics, 2019). Given demand (for electricity, transportation, steel, and heating), and policies, each FTT sub-model simulates technology diffusion in each of these sectors. The simulations start in 2013 and end in 2050. Any sub-model can be run either as a stand-alone model or in combination with the macro-econometric model E3ME[57] (Cambridge Econometrics, 2019) to create the much larger, integrated E3ME-FTT model (Mercure et al., 2014). In the integrated model, E3ME provides a disaggregated ('top-down') representation of the macroeconomy based on econometric relationships (Cambridge Econometrics, 2019) and FTT provides a ('bottom-up') description of technology dynamics. When running the full E3ME-FTT model, demand (for electricity, transportation, steel, and heating) is computed endogenously by E3ME (based on macroeconometric equations). The main difference when running FTT as a stand-alone model is that demand (e.g. for electricity) becomes an exogenous input. The run time is significantly longer for the full E3ME-FTT model than for FTT on its own: It takes about 30 seconds to run a scenario using the

[55] This centre no longer exist. A part of the research that previously took place in 4CMR now takes place in the Cambridge Centre for Environment, Energy and Natural resource Governance (C-EENRG).

[56] A sub-model for agriculture is also being developed this year.

[57] Previously called E3MG, also discussed in Chapter 4.

FTT power sub-model and about 30 minutes using E3ME-FTT[58]. E3ME-FTT can also be linked to the climate model GENIE in order to create a fully integrated IAM, called E3ME-FTT-GENIE (Mercure, Pollitt, Edwards, et al., 2018). Otherwise, emissions pathways from FTT and E3ME-FTT can be inserted manually into the GENIE model in order to obtain impacts of emissions e.g. on global temperature.

Mercure et al. describe E3ME-FTT-GENIE as "a fully descriptive, simulation-based integrated assessment model designed specifically to assess policies " (2018, p. 195). More specifically, the team who designed the full model write,

> "the modelling approach…is one of *simulation*. Each part of the E3ME-FTT-GENIE modelling framework attempts to represent *real world* [sic] relationships, in terms of accounting balances, physical interactions and human behaviour… The results from the model are *predictions of outcomes* based on empirical, behavioural and physical relationships observed in the past and the present." (Mercure, Pollitt, Edwards, et al., 2018, p. 196, my italics).

Thus, the term 'simulation' is used by Mercure et al. to denote a modelling approach that aims to represent real-world relationships as closely as possible. This use of the term 'simulation' is closely related to the etymology of the word: *simulation* stems from the Latin term *simulat-* which means copied or represented, and the verb *simulare*, which again stems from *similis*, means like (*Oxford Dictionary of English*, 2010). Thus, used in this way, 'to simulate', means 'to imitate' or 'to copy'.

For Mercure et al., the overarching goal of the modelling approach that is exemplified by the development of FTT, E3ME-FTT, and E3ME-FTT-GENIE is to "improve our ability to anticipate the effects of climate policies" compared to what can be done using "equilibrium and optimisation-based" (2016, p. 103) models. In meeting this goal, the ability of FTT (and associated models) to offer a "a more realistic modelling approach" (2016, p. 102) is seen as key. Thus, for Mercure et al., the reason for *simulating* as opposed to *optimising* is intimately related to the goal of providing better predictions of the impacts of climate policies. In fact, Mercure et al. do not only want to improve predictive capacity, but propose "a fundamental methodological shift in the modelling of policy" (2016, p. 102).

This is relevant because the majority of scenarios in the AR5 ensemble (at least 72% of the scenarios generated by the 14 most influential IAMs) are based on IAMs that rely on equilibrium or optimisation

[58] This difference in run time makes the sensitivity analysis conducted in Chapter 6 infeasible for the full model, which is why Chapter 6 runs FTT as a stand-alone model.

in some way (see Chapter 3). Among other things Chapter 4 suggested that optimality assumptions might bias the sign of mitigation costs in the AR5 ensemble. Chapters 2-4 of this thesis have all called for a diversity of approaches. Thus FTT, which by all accounts represents a "different" approach, should be worth looking into. According to Mercure et al. (2018), what they call the "descriptive" purpose of FTT distinguishes it from the vast majority of other energy system models, which are primarily based on optimisation[59].

In order to understand the context in which FTT was developed and why it is seen by Mercure et al. to offer a "more realistic modelling approach" (2016, p. 102) compared to optimisation-based models, it is useful to revisit some of the terms related to the distinction between optimisation and simulation in the climate policy and modelling literature.

5.2 Simulation and optimisation

The distinction between simulation and optimisation for policy models, of which IAMs can be considered a subset, is not new. Morgan and Henrion (1990) distinguish between models used for 'classical decision analysis' and 'predictive policy models'. In models used for classical decision analysis, the goal is to discover the optimal decision, D*, given a range of parameters, X. In classical

[59] Mercure et al. (2014; 2016; 2018), and some other authors (DeCarolis et al., 2017; Dodds et al., 2015), refer to optimisation models as "normative" models. I will not use this term in this way because it contradicts the way that the term normative is often used (including in AR5). For economists, the computation of a solution that minimises the cost (or some other variable) would be considered a "descriptive" exercise because the solution *describes* the pathway that – according to assumed parameter values and relationships – minimises the cost. Such statements are part of what economists normally call "descriptive" (or "positive") economics, i.e. statements about the world that can be considered objective or verifiable. "Normative" economics instead concerns itself with what *should* or *ought to* be. Normative statements rely on value judgments and cannot be tested or verified. According to this use of the terms, identifying the pathway that minimises cost would be descriptive, but advising policymakers that this is the best pathway (and thus the one to strive for) would be normative. In other words, whether or not a pathway that is computed by an IAM is *desirable* is a normative question, but the computation of the pathway itself is not. (This is not to say that descriptive statements cannot also involve value judgment. Both the choice of questions to analyse and the choices that have to be made in order to compute an answer often involve value judgments. This, however, is true for all of science and thus not particular to IAMs.) It is perhaps because outputs of optimisation models are often perceived to be, or in practice becomes recommendations that optimisation IAMs are sometimes described as "normative" models. That is, IAMs might be used to provide answers not only to how we *can* get from here to there, but how we *should* get from here to there. In order to avoid the confusion, the term optimisation model is used in this thesis to refer to models that rely on mathematical optimisation techniques.

decision theory the criterion used to determine the optimal decision is the maximum expected utility (MEU). Models used for 'classical decision analysis' can thus be depicted as

$$f(X, MEU) \to D*$$

In 'predictive policy models', the decision, D, is instead a model input. Usually, in these models, the output is the quantity of some criterion, such as net present value or utility, U. 'Predictive policy models' can be depicted as

$$f(X, D) \to U$$

A similar distinction is used by AR2 (IPCC, 1995) to distinguish between two different types of IAMs: 'policy evaluation models' (PEMs) and 'policy optimization models' (POMs). Whereas PEMs resemble 'predictive policy models', POMs resemble the models used for 'classical decision analysis'. More specifically, PEMs *evaluate* the effects of a given climate policy and POMs *identify* an efficient or cost-effective climate policy. PEMs are also described by AR2 as 'projection models' (IPCC, 1995). In general, PEMs are "process-based models that attempt to provide a thorough description of the complex, long-term dynamics of the biosphere-climate system" (IPCC, 1995, p. 372). While PEMs are rich in physical detail, they tend to include only a limited representation of the socioeconomic system. Typical outputs from PEMs include emissions, GHG concentration levels, temperature, sea level, land use, and physical impacts such as ecosystems at risk, coastal land area lost, and mortality rates (IPCC, 1995). POMs instead optimise over control variables to achieve specified policy goals (e.g. emissions reductions) in a way that maximises some quantity (e.g. utility or profit). POMs include both cost-benefit and cost-effectiveness models. In cost-benefit POMs, the policy goal might be the maximisation of welfare and the control variable might be the level of emissions reductions. In cost-effectiveness POMs, the policy goal is the minimisation of the cost of meeting a given emissions target and the control variable might be the carbon tax profile. The distinction between POMs and PEMs was also used in the third and fourth IPCC assessment reports (reflecting the IAM literature (e.g. Kann & Weyant, 2000; Löschel, 2002)) to describe key differences between IAMs. The distinction is not, however, used in AR5.

As already mentioned, the majority of the scenarios in the AR5 ensemble stem from IAMs that use optimisation in some form to compute transformation pathways. All the perfect foresight IAMs in AR5 (see Chapter 3), i.e. the optimal growth and ESOMs, can be considered POMs.[60] In these models, the quantity that is optimised is welfare or consumption (in optimal growth models) or total energy system costs (in ESOMs). In addition to this, optimisation is also central to CGE models. The latter models are (as noted in Chapter 3) similar to optimal growth models in that they compute a solution that corresponds to the optimising behaviour of agents[61], but CGE models add to this an optimisation over economic sectors (Wolf, Schutze, et al., 2016). CGE models are, however, frequently used in "simulation mode" for policy impact analysis, i.e. to predict the effects of policies rather than to optimise policies. Böhringer and Löschel (2006) provide a useful review of the use of CGE models for sustainability impact assessment and Scrieciu (2007) provides a critical response to this.

It is in the last category, the *Partial Equilibrium – Myopic* category, in Table 3.1 that we might find non-optimising IAMs. While neither Morgan and Henrion (1990) nor AR2 (1995) use the term *simulation*, other authors (Edenhofer et al., 2006; Nikas et al., 2019) have used this term to distinguish non-optimising models from optimising models[62]. Chapter 3 described simulation models (in line with Edenhofer et al. (2006)) as models that are based on algorithms that compute solutions at each point in time based on the state of the modelled system at the previous point in time. This stepwise computation process is based on assumptions about how the system behaves and responds to changes (e.g. as a result of agent behaviour). Thus, climate policies in simulation models are inputs, not outputs. Simulation models thus have a lot in common with the 'predictive policy models' and the PEMs described in this section. Some authors (e.g. Löschel, 2002) use the term simulation model synonymously with PEM.

Thus, the distinction between optimisation and simulation is closely related to the use of climate policies as inputs or outputs in the modelling: For optimisation models, in general, emissions targets are the inputs and climate policy (usually a carbon price profile) is the output (together with other outputs such as the cost of mitigation). For simulation models, in general, climate policy (carbon prices and often other policies such as subsidies and feed-in-tariffs) is the input and an emissions profile is the output

---

[60] However, these models can also be used in "simulation mode" to evaluate the impacts of policies. This will be discussed in section 5.3.

[61] It should be noted that the lack of foresight (i.e. myopic behaviour) in the CGE IAMs in AR5 generally means that the solutions computed by these models are not globally optimal. Still, agents in these models make their decisions by optimising (with limited foresight) at each point in time. See Keppo and Strubegger (2010) for a demonstration of the impacts of lack of foresight in energy system models.

[62] The term 'simulation' is sometimes used synonymously with 'computer simulation', which simply means running a computer program. If used in this way, simulation is not opposed to optimisation, because any computer program might also be optimising, and optimisation IAMs are always run on computers. 'Simulation' in this thesis refers to the solution algorithm that is used in an IAM, as explained in this chapter.

(together with other outputs such as the deployment of different technologies). Thus, the distinction between simulation and optimisation is intimately related to the distinction between "backcasting", in which the end goal (such as an emissions target) is set, and "forecasting", in which the model is used to predict where we are headed (see Robinson (1982) for an early discussion of backcasting versus forecasting). Optimisation models can, however (as already noted), also be used in "simulation mode" to simulate the impacts of policies (i.e. optimisation models can be used as PEMs). For simulation models, however, the opposite is generally not the case: simulation models cannot be used to compute optimal policies[63].

Because FTT was designed to offer an improved representation of energy system dynamics relative to ESOMs, this chapter focuses on ESOMs (as opposed to optimal growth and CGE models). The next section explains how ESOMs, from their inception, have been used both to identify optimal policies (in "optimisation mode") and to predict the evolution of energy systems (in "simulation mode").

5.3 Optimisation as prediction

Well known ESOMs such as MARKAL and MESSAGE[64] (and their forerunners, see Figure 2.3) have been used since the 1970s to compute technology pathways that minimise overall energy system costs subject to emissions constraints, based on linear programming techniques. These models were designed to analyse the transition away from fossil fuels. If we look at their history, we find that they have been seen both as tools for the identification of optimal solutions – which is useful for energy planning – and as tools that describe the actual evolution of the energy system.

This dual role can be seen, for instance, in the history of the MESSAGE energy system optimisation model. Despite the Hafele/Manne model being the "grandfather" of MESSAGE and the fact that the two models share the same model structure, the interpretation of the pathways generated by the two models differed. Optimisation in the Hafele/Manne model was seen to offer a way of representing the real-world dynamics of the energy system (Häfele & Manne, 1974). That is, the energy system was

---

[63] Or rather, it would be very cumbersome to do so, as one would have to go through every possible option.

[64] Much like FTT has been integrated with the macro-econometric model E3ME, MESSAGE and MARKAL have been integrated with the macroeconomic model MACRO. All three models can be run either as stand-alone models or as part of larger integrated models. The integrated models are sometimes referred to as 'hybrid' models because they combine detailed energy system models with macroeconomic models. The terms 'bottom-up' and 'top-down' are also used to refer to energy system IAMs and macroeconomic IAMs respectively. Although this distinction has been used widely to explain differences between IAMs in the past, it has recently become much less relevant because many IAMs are now hybrid.

seen to evolve roughly along the lines of the least-cost pathway. Optimisation in MESSAGE, instead, was seen to reflect the centralised structure of decision making in the energy system (Agnew et al., 1979b). More specifically, Agnew et al. write, “the structure of the energy supply system is the result of a few crucial and far reaching decisions taken at governmental or supra-national level” rather than “the perfect market of the classical economists – which assumes a large number of independent actors” (1979b, p. 5). MESSAGE was thus designed to answer questions such as: “What is the optimal timing for the implementation of new energy supply technologies?”, and “What constraints does the environment impose on the “optimal” energy strategy?” (Agnew et al., 1979b, pp. vii–viii). In other words, MESSAGE was conceived of by Agnew et al. (1979b) as a planning instrument to be used by national and international decision makers in order to achieve an optimal allocation of resources (given energy demand and environmental constraints), not as a tool to predict the evolution of the energy system.

EFOM (the forerunner to MARKAL), which was developed around the same time as the Hafele/Manne model, was also seen to describe the evolution of the energy system. In describing EFOM, Finon writes that “our choice was directed towards a tool for economic calculation, which immediately implies *the representation of the real world* by means of the relations between economic agents or operations expressed quantitatively in terms of physical flows and internal prices. A sectorial optimisation model has thus been worked out, integrating the internal interdependences (relations between economic operations) and external interdependences (constraints or exogenous parameters)” (Finon, 1974, p. 138 my italics). In short, both EFOM and the Hafele/Manne model used optimisation to describe the real-world evolution of the energy system, not just to identify the least-cost solution, which might be implemented by a centralized planner.

Thus, optimisation in ESOMs played a dual role from the beginning. On the one hand optimisation models were seen to represent the actual behaviour of systems, and on the other hand it was seen as a tool to help energy planners identify and enact the most rational decisions[65] (for a system that might not otherwise behave rationally). Both perspectives are still evident today and explanations of how ESOMs describe the real-world evolution of the energy system can be found in the model documentation also for more recent ESOMs (e.g. Loulou and Labriet (2008)). The next section explains the assumptions that are either explicitly or implicitly made when one uses ESOMs to predict the evolution of energy systems including its response to policies.

[65] It should be noted that if one believes that the central planner will make the cost optimal decisions, then the energy system *will* evolve according to least cost. Still, the reason for why it would do so differs.

5.3.1 Limitations to neoclassical economics assumptions

In short, the assertion that optimisation can be used to represent the real-world response of the energy system to climate policies is based on a number of well-known neoclassical economics assumptions: If we assume perfect (energy) markets and fully rational agents with full information, energy supply and demand will meet (equilibrium will be achieved) in such a way that consumer and producer surplus is maximised (Mas-Colell et al., 1995). Loulou and Labriet's (2008) description of TIMES/MARKAL, for example, is based on these assumptions. An equilibrium, they write, is the state in which prices and quantities are such that no consumer would want to purchase less than quantity q* at price p*, and no producer would want to produce more than quantity q* at price p*. Under these conditions, total surplus is maximized. If we believe that the assumptions of perfect markets and fully rational agents provide a sufficiently good approximation of the real world, we might use maximisation of total surplus as a way of deriving the real-world response of the energy system to changes, such as those caused by climate policies. Several examples of such usage of ESOMs can be found in the literature.

Hu and Hobbs (2010), for example, use MARKAL to "simulate[] the operation of a competitive market with zero price elasticity for energy services". Altamirano et al. (2008) couple the MARKAL model for Switzerland with a CGE model (GEMINI-E3) in order to simulate the impacts of suggested climate policies. MARKAL is also used e.g. on EU level to assess the implications (i.e. the response of the energy system) of selected policy instruments including white and green certificates and emissions trading (Mundaca & Santi, 2004). Babaee et al. (2014) use TIMES to explore how the deployment of electric vehicles in the future is affected by natural gas prices, oil prices, battery costs, renewable portfolio standards and CO2 prices. Pye et al. (2015) use another ESOM, ESME[66], to predict the impacts of policies on emissions.

The issue when it comes to the use of ESOMs in "simulation mode" to describe the evolution of real-world energy systems, including the impacts of policies, has to do with the limitations to the above assumptions. The assumptions of fully rational and perfectly informed optimising agents have long been criticised in the economics literature (Veblen, 1898). Important contributions from within economics, such as the Sonnenchein-Mantel-Debreu theorem from the 1970s, and the Greenwald-Stiglitz theorem from the 1980s, have even led to the questioning of the existence of equilibria and the view that market failure is not the exception, but the norm in economics (Mas-Colell et al., 1995). More recently, the economics discipline has also evolved to incorporate new areas of research, such as behavioural economics, which is set out specifically to study behaviour that deviates from the purely rational. It is increasingly accepted that neoclassical assumptions of perfect markets and fully rational

[66] The Energy Systems Modelling Environment (ESME) model is mathematically similar to MARKAL-TIMES.

agents provide an overly simplistic picture of how the real economy works (see e.g. Chang (2014)). In particular since the great financial crisis, the economics discipline has been under increasing scrutiny (A. Turner, 2012). For all these reasons, assuming perfect markets and rational behaviour could be viewed, not only as unrealistic, but outdated. Although few examples exist of studies that attempt to test whether ESOMs provide a good description of real-world energy systems, one such study (Trutnevyte, 2016) suggests that this is not the case.

It is in this context that FTT, according to Mercure et al. provides "a more realistic modelling approach" (Mercure, Pollitt, et al., 2016). ESOMs are generally characterised by minimisation of overall energy system cost using linear programming and ESOMs generally assume perfect foresight. According to DeCarolis et al. (2017, p. 188) "in its most basic form, an ESOM makes optimal technology investment and utilization decisions based on differences in the relative cost of competing technologies, thermodynamic performance limits, fixed end-use demands, and constraints that reflect known physical resource limits or policy objectives. The associated model-based results provide a prescription that indicates what should happen if a single rational economic decision maker acts from a social planning perspective to minimize cost". It is important to recognise, however, that ESOMs have also been modified in many ways in attempts to increase realism. Additional features include ETL, lumpy investments (i.e. discrete sizes of certain technologies) and hurdle rates, to name a few (see DeCarolis et al. (2017) for more examples). The point here is not to criticise ESOMs, however, but to understand the motivation behind the construction of FTT and the ways in which FTT was seen, by those who built it, to offer an improved representation of the response of the energy system to climate policies.

The main distinguishing features of FTT are described in the next section. Before so doing, the following point should be noted. Although neoclassical assumptions of rational behaviour and perfect foresight have been widely criticised and are viewed by many as overly simplistic, all models represent simplifications of the real world and all model outputs depend on assumptions. Unless we can show that FTT is based on more accurate or better justified assumptions, we cannot claim that FTT provides more accurate descriptions of energy systems and thus more accurate predictions than ESOMs. The next section presents the key features that distinguishes FTT from ESOMs before explaining why technology deployment assumptions play a crucial role in both FTT and in ESOMs.

5.4 The importance of technology deployment rates

FTT was designed with two specific features in order to improve the realism of the energy system pathways it depicts:

(i) A more fine-grained representation of policies (including carbon prices, subsidies, regulations, feed-in-tariffs, and kick-start policies (Mercure, Pollitt, Edwards, et al., 2018)).

(ii) An endogenous derivation of technology deployment rates based on diffusion theory (Mercure, 2015).

So far, FTT publications have focused on the first feature. E3ME-FTT has, for instance, been used to demonstrate how baskets of policies can be used to achieve desired emissions reductions (Mercure, Pollitt, Edwards, et al., 2018). But the second feature is equally crucial for the predictions of policy impacts. Diffusion theory has, according to Mercure (2015, p. 3), "yet to be even considered in large scale mainstream models such as those for energy system modelling and related energy policy analysis". The endogenous derivation of deployment rates is thus a unique feature of FTT seen by Mercure (2015) to offer improved and theoretically grounded predictions of the impacts of policies on technological change.

Given the limited emphasis on this feature in FTT publications so far, however, a better understanding of what the diffusion theory that underpins FTT implies for the reliability of predictions is needed. In order to begin to assess the accuracy of FTT predictions, Chapter 6 conducts the first comprehensive sensitivity analysis focused specifically on the assumptions that determine the endogenous derivation of technology deployment in FTT. Before then, we take a brief look at how technology deployment is determined in ESOMs, which represents the main contender to the approach taken in FTT.

5.4.1 Exogenous technology deployment rates

ESOMs are solved using linear programming (also called linear optimisation) techniques. A well-known feature of these techniques is that that small changes in parameter values or variables can result in dramatic changes to outputs, a feature that is sometimes referred to as 'penny-switching' or 'flip-flopping' behaviour (Keepin & Wynne, 1984; Wilson et al., 2013). The reason for this is that in linear programs, the optimal solution is always found in one of the corners (so-called "corner solutions"). This means that an infinitesimal change in parameter values can lead the hyperplane that defines the feasible set of solutions to change its tilt from one direction to another, causing the optimal solution to move from one corner to another (see Keepin & Wynne (1984) for a graphical depiction of this).

For ESOMs, this means that tiny changes, in for instance technology costs or learning assumptions, can lead to sudden changes in technology choices. In order to avoid situations in which all investments are suddenly moved from one technology to another – which would be considered unrealistic[67] – ESOMs are constrained by maximum deployment rates, which are provided as input assumptions. Such

[67] I thank Will McDowall for an insightful discussion regarding this.

constraints are present in all of the 14 most influential IAMs in AR5 analysed in Chapter 2 that are based on energy system optimisation[68].

The issue with maximum deployment constraints is that they are often binding – meaning that the rate of deployment (of some technologies) will be defined by them – but based on limited empirical evidence. A sensitivity analysis of the widely used MARKAL model, for example, showed that the deployment constraints for electricity generating technologies had significant impacts on results (Johnson et al., 2006). The issue becomes even more pertinent when learning-by-doing is included. The TIMES manual states that when endogenous learning is switched on, the model might produce "unrealistically large early investments in some learning technologies" (Richard Loulou & Labriet, 2008, p. 30). For this reason, modelers "impose additional constraints to control the penetration of learning technologies…reflecting what is deemed realistic. These upper bounds play a determining role in the solution of the problem, and it is most often observed that the capacity of a learning technology is either equal to 0 or to the upper bound. This last observation indicates that the selection of upper bounds by the modeler is *the predominant factor in controlling the penetration of successful learning technologies*" (Richard Loulou & Labriet, 2008, p. 31 my italics). For this reason, Loulou and Labriet question whether it is "worth-while for the modeler to go to the trouble of modeling endogenous learning (with all the attendant computational burdens) when the results are to a large extent conditioned by exogenous upper bounds" (Richard Loulou & Labriet, 2008, p. 31). If we had solid information on which to base the maximum deployment constraints, this dependency would not be such a problem. However, deployment constraints appear to be based on limited empirical evidence. According to Wilson et al. "constraints on technological growth is significantly less well substantiated and documented compared to the techno-economic parameterization…systems interactions…and learning processes" (2013, p. 383). What this means is that technology deployment in pathways generated by ESOMs might be determined entirely by assumptions that are based on weak (or even missing) empirical data. This poses a threat to the predictive value of ESOMs. The few studies on this that exist indicate that exogenous deployment constraints imply a conservative picture of change (Wilson et al., 2013).

---

[68] In fact, constraints on the deployment of technologies are present in all the *Perfect foresight* of the 14 IAMs (which all rely on global optimisation). That is REMIND, MESSAGE-MACRO, WITCH, MERGE-ETL, MERGE, DNE21+, TIAM-World, and BET all have built-in assumptions regarding either the maximum percentage growth of technologies from one year to the next (MESSAGE, TIAM-World, BET) or adjustment costs that makes it increasingly expensive to deploy technologies faster (REMIND, MACRO). This information is obtained from the model documentation and from two papers (Bauer et al., 2018; Pietzcker et al., 2017).

5.5 Conclusion

This chapter has shown that not only neoclassical assumption, which have been widely criticised, but also deployment rate constraints, which appear to be based on limited empirical evidence, challenge the realism of transformation pathways generated by ESOMs. In short, transformation pathways generated by ESOMs cannot be expected to describe the real-world evolution of the energy system if (i) neoclassical assumptions of perfect markets and fully rational agents do not hold (sufficiently), or if (ii) exogenous deployment constraints do not accurately capture real-world constraints.

According to Mercure et al. (2014; 2016), the simulation approach taken in FTT (and E3ME-FTT) distinguishes it from the majority of IAMs, which mostly rely on optimisation. In particular, the endogenous derivation of deployment constraints is seen by Mercure et al. to make FTT unique compared to all other energy system models. Together, these features are seen to offer a more realistic depiction of the impacts of policies on future technology transformations, grounded in the theory of technology diffusion.

The fact that technology deployment rates are determined endogenously in FTT, however, does not imply that they do not depend on exogenous assumptions. All model outputs depend on exogenous assumptions (this is, in many ways, what defines a model). The difference is that, while it is clear what assumptions determine deployment rates in ESOMs, the answer is more complex for FTT.

In order to determine whether FTT provides more accurate predictions of technology evolution compared to ESOMs, we need to determine both whether the endogenous description of technology deployment represents an accurate description of technology deployment in the real world and whether the parametric assumptions that determine technology deployment in FTT are better justified than the maximum deployment constraints that are used in ESOMs. The next chapter takes the first step in answering the latter question by conducting the first global sensitivity analysis of the core FTT equation.

6 Assessing FTT predictions

In order to begin to assess the uncertainty of FTT predictions, this chapter conducts the first global sensitivity analysis of the FTT model. In order to render the task feasible, the analysis is applied to the power sector sub-model, FTT:Power, only. The power sector is, because of its contribution to global GHG emissions and because it is one of the cheaper sectors to decarbonise, seen as one of the most important sectors to decarbonise first (IPCC, 2014a). In 2018, total CO_2 emissions were estimated at 41.5GtCO_2 (Quéré et al., 2018). Of these, 13 GtCO_2 came from the power sector alone (IEA, 2019).

Power sector CO_2 emissions can be reduced in three different ways: by reducing end-use electricity demand, by increasing efficiency of production and consumption, and by substituting low-carbon technologies[69] for unabated fossil fuel technologies. Global electricity demand is expected to increase (due in particular to the increase from developing countries), however, and efficiency can only be improved up to a certain limit. It therefore follows that, in order for power sector emissions to reach net zero, electricity generation can no longer be based on unabated fossil fuel technologies. Thus, while reducing demand and increasing efficiency represent crucial avenues for emissions reductions, these measures will not be sufficient to meet the Paris target. As long as there is demand for electricity, the rate of emissions reductions in the power sector will depend crucially on the rate at which low-carbon technologies replace unabated fossil fuel technologies. FTT:Power was designed specifically to predict the impacts of policies on technology substitution in the power sector.

So far, only best guess predictions from FTT have been published in the literature (Mercure, 2012; Mercure et al., 2014; Mercure, Pollitt, Viñuales, et al., 2018). In order to provide a better picture of the uncertainty of FTT predictions, which is a prerequisite for evaluating their accuracy, this chapter conducts a global sensitivity analysis of FTT:Power predictions. Although the analysis does not include all FTT:Power parameters, it does include all the parameters that define the core equation, which is shared between all FTT models. The goals of the analysis are to i) provide a first conservative estimate of the uncertainty of FTT:Power predictions and, ii) to identify the parameters in the FTT core equation that have the biggest influence on FTT predictions. In order to provide an initial (conservative) estimate of the uncertainty, uniform and independent distributions based on varying default parameter values by ±50% are assumed. The results should only be interpreted as a first order, conservative, estimate of the full uncertainty of FTT:Power predictions: a full uncertainty assessment would need to take into account not only the uncertainty of all parameters but also structural uncertainty.

[69] Low-carbon technologies include renewables, nuclear, fossil fuel technologies with carbon capture and storage (CCS), and negative emission technologies (NETs).

The results of the sensitivity analysis conducted in this chapter show that, although FTT:Power predictions indicate that policies are necessary to reduce emissions sufficiently to meet the Paris target, the impacts of policies depend crucially on a scaling parameter whose correct value is deeply uncertain. This finding challenges the reliability of best guess FTT:Power predictions. Given the importance of the uncertainty of the emissions reductions caused by climate policies, this chapter argues, providing a range of output values is more appropriate than providing single best guess predictions.

Section 6.1 presents the best guess predictions of the impacts of policies aimed at limiting global warming to 1.5˚C and 2˚C generated by FTT:Power. Section 6.2 explains the logic behind FTT:Power, the core equations, and the parameters that will be included in the sensitivity analysis. Section 6.3 reviews previous sensitivity analyses of FTT and uncertainty assessment of IAMs in general. Section 6.4 explains the methods used for the sensitivity analysis – Monte Carlo analysis and Latin Hypercube Sampling – and discusses the use of independent and uniform distributions to define the ranges of parameter values. Section 6.5 presents the results. Policy implications are discussed in section 6.6 and concluding remarks are provided in section 6.7.

6.1 Best guess predictions

The modelling team behind E3ME-FTT-GENIE has used the model to identify two different sets of policies that, if implemented, will ensure that the Paris target is met. The '2C Policies'[70] will, according to the model, limit warming to below 2˚C, and the '1.5C Policies'[71] will limit warming to below 1.5˚C. These policy sets thus generate two distinct mitigation scenarios. In addition to this, there is also a baseline scenario that corresponds to what happens when 'No Policies' are implemented.

Optimisation models can take climate targets as inputs and compute optimal policies as outputs. In E3ME-FTT policies are instead given as inputs. This means that the identification of policy sets that meet specified climate targets has to be done by a process of trial and error: Policies deemed plausible are added one by one and resulting emissions are recorded. If emissions are too high, policy measures are tightened. If emissions are too low, policies are relaxed. This is repeated until the desired target is achieved. A detailed example of the process of trial and error for identifying policies in the power sector can be found in Mercure et al. (2014).

[70] See (Mercure, Pollitt, Edwards, et al., 2018).

[71] No scenarios based on the 1.5C policy set have yet been published. At the time of writing, an E3ME-FTT-GENIE paper presenting this policy set was considered for re-submission.

The policies that are identified using E3ME-FTT (or FTT) are not unique. That is, different combinations of policies could have equally well been used to achieve the same climate target. FTT:Power, for example, takes five different policies as inputs: carbon prices, subsidies, feed-in tariffs (FiTs), regulations, and kick-start policies. This means that identical emissions pathways could be generated by for instance decreasing renewables subsidies and increasing the carbon price. Because E3ME-FTT is not an optimisation model, the policy sets that are identified do not represent the optimal policy sets. The policies chosen by the modelers are instead based on their judgments of what is deemed realistic[72].

The impact of '2C Policies', '1.5C Policies', and 'No Policies' on the deployment of the 24 technologies that are represented by FTT:Power is shown in figure 6.1. These results are obtained by using the three policy sets as inputs to FTT:Power and by taking the electricity demand from the corresponding scenarios computed by the full model[73]. The use of coal for electricity generation is significantly reduced in the two mitigation scenarios. This is not surprising, given the carbon intensity of electricity produced from coal. The 1.5C and 2C policies also lead to a reduction in the use of gas (CCGT). The reduction in the use of coal and gas is compensated for by the deployment of solar PV and wind (primarily onshore). We also see an increase in the use of nuclear, solid biomass, and biomass (BIGCC) with carbon capture and storage (CCS) as a result of the two policy sets. As expected, the effects are more pronounced in the 1.5C scenario than in the 2C scenario. Especially the deployment of BIGCC + CCS, which is a negative emissions technology (NET), is significantly higher in the 1.5C scenario than in the 2C scenario[74].

---

[72] Personal communication with Jean-Francois Mercure and Hector Pollitt (Cambridge Econometrics). Cambridge Econometrics, which is responsible for E3ME-FTT, has many years of experience working with policy analysts, in particular on EU level. This experience is partly what shapes the judgments regarding the realism of policy baskets.

[73] In the full E3ME-FTT model, E3ME and FTT:Power are solved iteratively until the (regional) electricity demands computed by E3ME and the (regional) electricity prices computed by FTT:Power are consistent. Everything else being the same, an increase in the price of electricity leads to a reduction in electricity demand. The projections in Figure 6.1, however, were obtained by running FTT:Power as a stand-alone model. In order to re-produce the power sector predictions caused by the policy sets in the full model using FTT:Power alone, the electricity demand that was computed endogenously with the full model is used as an input to FTT:Power.

[74] The use of NETs in scenarios that reach the Paris target have been heavily debated in the last few years (Anderson & Peters, 2016; Fuss et al., 2014). Even though the use of NETs is limited in the FTT projections, it is worth keeping in mind that FTT projections only extend to 2050. The use of NETs in scenarios increase primarily in the latter half of the century.

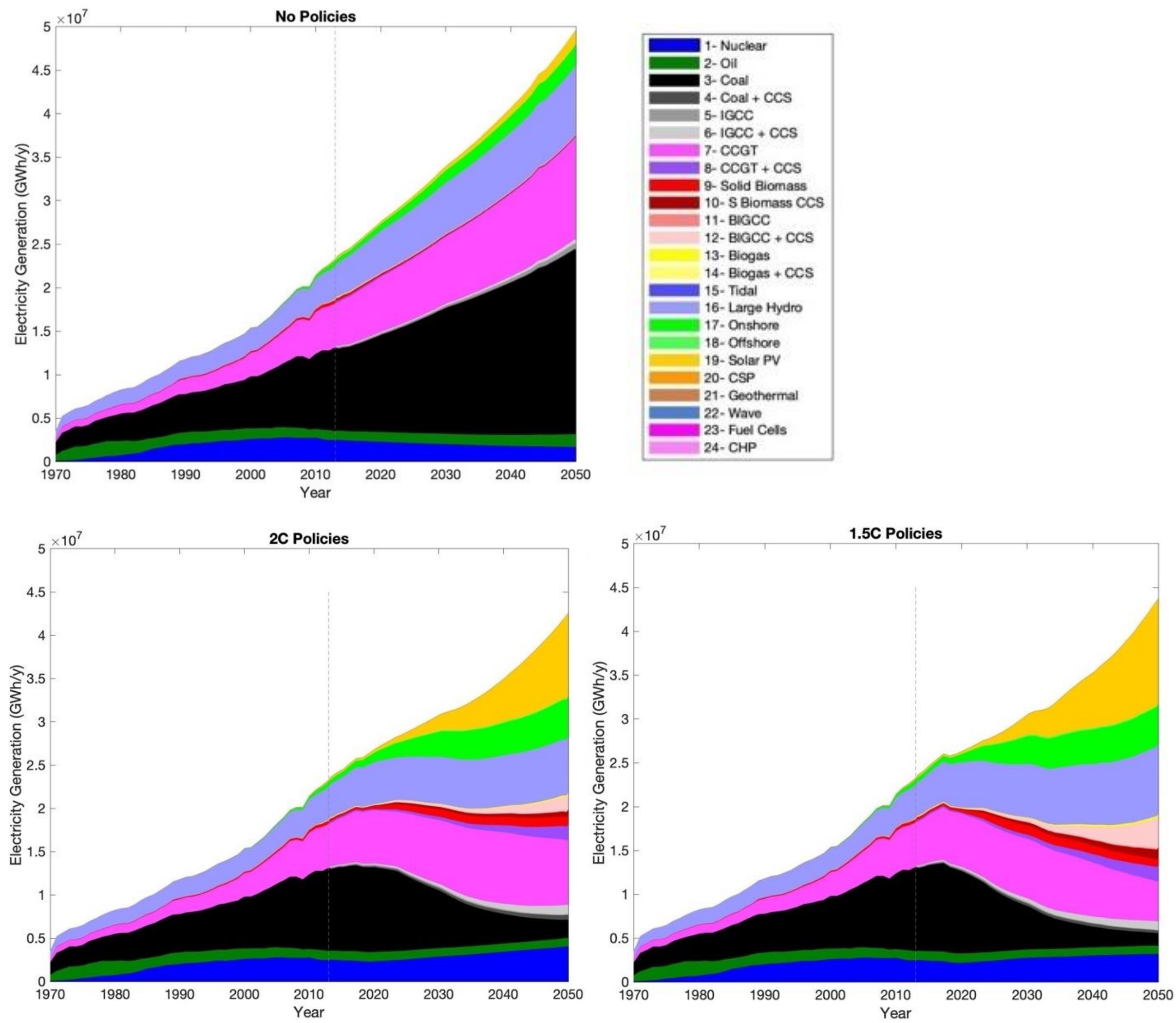
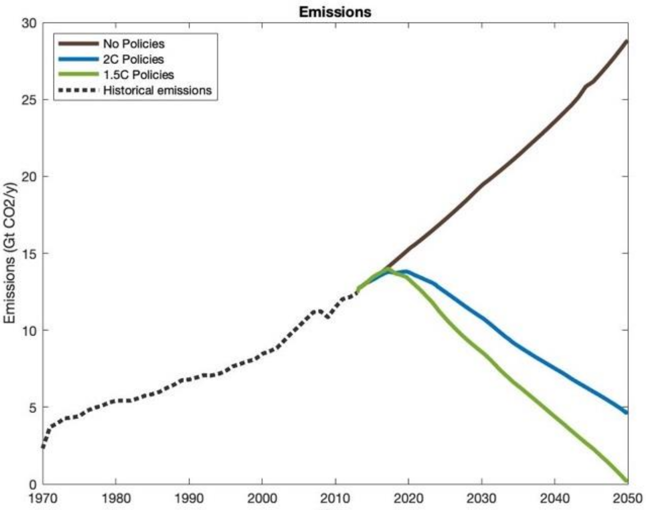


Figure 6.1 Predictions ('best guess') of technology deployment in the power sector caused by No Policies (top), 2C Policies (middle), and 1.5C Policies (bottom). The dashed line shows the start date of the simulations (2013). The figure is produced by the author using FTT:Power with electricity demands taken from the corresponding runs in the full E3ME-FTT model.

Figure 6.2 shows the resulting power sector emissions and cumulative emissions. The figure shows rapid emissions reduction in the two mitigation scenarios. The cumulative emissions in the 2C scenario is roughly half of the emissions in the baseline scenario. The difference in cumulative emissions between the 2C and 1.5C mitigation scenarios (80 $GtCO_2$), however, corresponds to only six years of power sector emissions at the current rate (13 $GtCO_2$/yr).



Scenario	Cumulative emissions ($GtCO_2$) 2013-2050
No Policies	749
2C	372
1.5C	292

Figure 6.2 Predictions ('best guess') of power sector emissions caused by No Policies, 2C Policies, and 1.5C Policies in FTT:Power.

So far, the emphasis in FTT and E3ME-FTT publications has been on using the distinguishing features of these models – in particular the simulation approach and the fine-grained representation of policies – to either identify policies that will lead to the desired emissions reductions, or to predict the impacts of given policies, with an emphasis on realism. So far, this has been done (e.g. by Mercure (2012) and Mercure et al. (2014; 2018)) using 'best guess' parameter values. In order to understand how FTT:Power computes the impacts of policies on technology deployment and emissions in the power sector (shown in figures 6.1 and 6.2), and the parameters involved, the next section presents the core FTT:Power equations.

6.2 FTT core equation

FTT is best described as a set of technology diffusion models. Diffusion represents the gradual adoption of innovations by firms and individuals. It is the third stage in the process of technological change as described by Schumpeter (1942), the first stage being invention, and the second stage being innovation. Diffusion follows two key observed characteristics. First, it is never instantaneous (Rogers, 1962). Second, it follows S-shaped curves – a characteristic that has been confirmed many times in the case of the evolution of competing energy technologies (e.g. Grübler et al., 1999; Marchetti & Nakicenovic, 1979).

All the FTT models are based on the same mathematical structure. That is, a set of coupled logistic differential equations based on the Lotka-Volterra family - also known as the predator-prey equations. This well-known set of equations is commonly used to describe the dynamics of biological systems. In FTT, it is used to represent the nature of technology diffusion. According to Mercure (2012), FTT:Power is the first power sector model to be based on this set of equations. Mercure (2012) also shows that it gives rise to classic S-shaped technology transitions, similar to those observed in history. This section describes how the set of coupled logistic differential equations that make up the core of FTT:Power is derived.

6.2.1 Species analogy

Mercure (2011) introduces FTT for the first time using the following analogy: Imagine a biological system made up of one species of birds only, which can nest only in a certain type of tree. If the species has fertility factor b, the rate of change of the fraction of trees occupied by it, $N(t)$, is given by the Verhulst equation,

$$\frac{dN(t)}{dt} = bN(t)\big(1 - N(t)\big). \tag{1}$$

The solution to this equation is the logistic function

$$N(t) = \frac{e^{bt}}{1 + e^{bt}}, \tag{2}$$

which exhibits exponential growth at low levels of N and saturation at high levels of N, as shown in figure 6.3. Over time (provided $N(0) > 0$), the solution converges to $N = 1$. The time that it takes for N to reach 1 is determined by the parameter b.

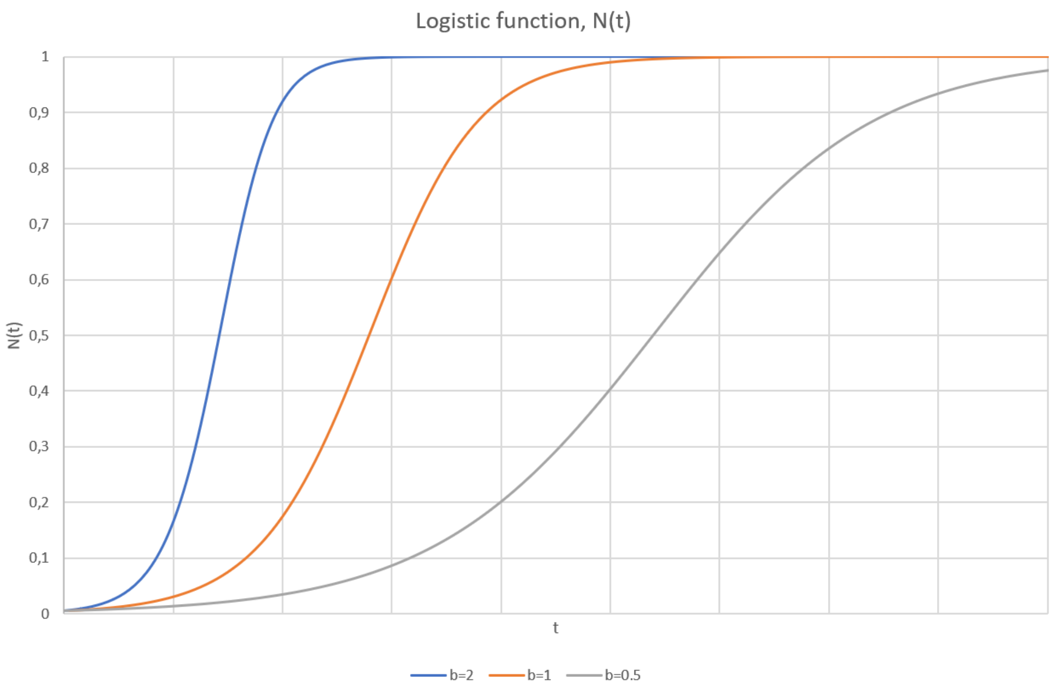


Figure 6.3 The logistic function for different values of b.

Mercure (2011) continues this analogy by imagining instead that there are two species of birds, N_1 and N_2, competing for space in the same type of trees. The parameter b in this case represents the ability of species one to capture nesting space *at the expense of* species two[75]. In this case, the rate of change in the fraction of space occupied by species one is given by:

$$\frac{dN_1(t)}{dt} = bN_1(t)N_2(t) = bN_1(t)\big(1 - N_1(t)\big) \tag{3}$$

As we can see, the expression for N_1 in equation (3) ends up being the same as the expression for N in equation (1) (as expected, given that N_2 simply represents what was previously the remaining empty trees). For positive b (and $N_1 > 0$), species 1 will eventually occupy all the space (for negative b, species 2 will).

6.2.2 Shares equation

In FTT, instead of species, there are technologies, and instead of nesting space or trees there are market shares. *The fundamental assumption is that technologies within a specific sector (such as power or transportation) compete for market shares in a way that is analogous to the way in which species compete for resources.* This assumption captures two aspects of technology systems dynamics: first, the role of competition between different technologies, and second, the idea that more established technologies can increase their market shares faster than nascent technologies. This set of assumptions are shared among all FTT models. The specific implementation, however, varies between FTT models.

In FTT:Power, the market share of each electricity generating technology, $S_i(t)$ is defined as

$$S_i(t) = \frac{U_i(t)}{U_{tot}(t)}, \tag{4}$$

where $U_i(t)$ denotes the total installed capacity (in GW) of technology i, and $U_{tot}(t)$ is the sum of $U_i(t)$ over all technologies, $U_{tot}(t) = \sum_i U_i(t)$. The shares evolve, as will be shown, according to a generalised version of equation (3), called the *shares equation*. It is the shares equation that (endogenously) determines the rates of deployment for technologies in FTT. The parameters analogous to the parameter b in equation (3), i.e. the technology substitution parameters, determine the ability of one technology to take market shares from another technology. In FTT:Power, the substitution parameters depend on two aspects. First, the costs of different technologies, which is seen to determine

[75] The implicit assumption here is that all the nesting space is always occupied. If it is not occupied by species 1 it is occupied by species 2.

investor preferences. And second, characteristic time constants, which determine potential turnover rates.

Investor preferences

The role of costs in defining the substitution parameters has to do with how FTT:Power models technology diffusion by "mimicking the decision-making of investors" (Mercure, 2011, p. 7). Essentially, investors' preferences for electricity generating technologies are assumed to depend on the cost of generating electricity using that technology, that is, the levelized cost of electricity (LCOE). The LCOE for technology i is given by

$$LCOE_i(t) = \frac{\sum_{t=0}^{\tau_i} \frac{TI_i(t) + OM_i(t) + FC_i(t) + CC_i(t)}{(1+r)^t}}{\sum_{t=0}^{\tau_i} \frac{EP_i(t)}{(1+r)^t}}, \tag{5}$$

where $TI_i(t)$ denotes the investment cost, $OM_i(t)$ the operation and maintenance cost, $FC_i(t)$ the fuel cost, $CC_i(t)$ the carbon cost for technology i, r the investor discount rate (hurdle rate), and $EP_i(t)$ the amount of electricity produced per year. Because of local variations, the investment costs ($TI_i(t)$), operation and maintenance costs ($OM_i(t)$), and fuel costs ($FC_i(t)$) are expressed as distributions rather than as point estimates. This means that the LCOEs themselves are also distributions. Data for investment, operation and maintenance, and fuel costs in the most recent version of FTT (v7) are obtained from the International Energy Agency's (IEAs) *Projected Costs of Generating Electricity 2015* (IEA & NEA, 2015). (The values of these parameters and their standard deviations are also listed in Appendix E.) The carbon cost ($CC_i(t)$) is a variable that represents the carbon price, which is determined by policies in FTT:Power (more on policy variables below).

The inclusion of learning-by-doing, is also a key feature of FTT:Power. This is incorporated via experience curves, which means that the cost of a technology decreases with cumulative installed capacity. In FTT:Power, it is the investment cost that is reduced via learning. This is done according to

$$TI_i(t) = TI_{0,i}\left(\frac{W_i(t)}{W_{0,i}}\right)^{-\beta_i}, \tag{6}$$

where $TI_{0,i}$ denotes the investment cost when cumulative installed capacity is $W_{0,i}$, and β_i is the learning coefficient for technology i. The total installed cumulative capacity, W_i also takes into account installed capacity of related technologies according to a learning spillover matrix (see Mercure, 2012). The learning rate, which describes the fractional reduction in cost for each doubling of cumulative capacity (Rubin et al., 2015), is then given by

$$LR_i = 1 - 2^{-\beta_i} \tag{7}$$

The fact that technology costs are represented by distributions rather than point estimates in FTT:Power means that even if a technology on average is more expensive than another technology, there can still be cases in which the opposite is true. This is illustrated in figure 6.4, which shows that, as long as there is sufficient overlap between the cost distributions, there are cases in which the (on average) more expensive technology is cheaper than the (on average) cheaper technology.

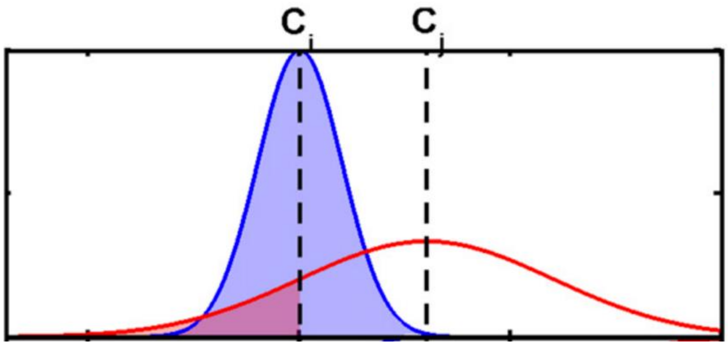


Figure 6.4 Probability distributions for the cost of two different technologies (with average costs Ci and Cj). Source: Mercure (2012).

This can be used to derive the cumulative probability distributions, $F_i(C, C_i)$, which give the probability that the cost of technology i, with average cost C_i, will be less than C. Based on this, investor preferences for electricity generating technologies in FTT:Power are assumed to be proportional to the probability that the cost of a unit of technology i is less than the median value of the cost of technology j. This is given by

$$F_{ij} = F_i(C_j, C_i) = \frac{1}{1 + \exp\left(\frac{C_j - C_i}{\sigma_{ji}}\right)}, \tag{8}$$

where $\sigma_{ji} = \sqrt{\sigma_i^2 + \sigma_j^2}$, and σ_j and σ_i are the standard deviations for the LCOE distributions for i and j, and $F_{ij} + F_{ji} = 1$ (Mercure, Pollitt, Edwards, et al., 2018). The shaded area in figure 6.4 corresponds to $F_{ji} = F_j(C_i, C_j)$. This way of representing investor preferences ensures smoother transitions. (If costs instead were represented as point estimates, investor preferences would suddenly switch when the LCOE of one technology, due for instance to learning, drops below the cost of another.)

Characteristic time constants

The substitution parameters also depend on technology characteristics that in FTT:Power determine potential turnover rates (Mercure, 2015). These are: the rate of retirement of technology j, $1/\tau_j$ where τ_j is the lifetime of technology j, and the rate at which units of i can be built, $1/t_i$ where t_i is the build time for technology i.

Based on investor preferences, and retirement and build rates, Mercure (Mercure, 2012) defines the (gross) flow of market shares from technology j to technology i to be proportional to

$$\frac{dS_{i\to j}}{dt} \propto F_{ij}\frac{S_i}{t_i}\frac{S_j}{\tau_j}, \tag{9}$$

and the (gross) flow of market shares from technology i to technology j to be proportional to

$$\frac{dS_{j\to i}}{dt} \propto F_{ji}\frac{S_j}{t_j}\frac{S_i}{\tau_i}. \tag{10}$$

The flows in equation (9) and (10) are independent. This means that the net flow of shares from i to j is given by

$$\frac{dS_{ij}}{dt} = KS_iS_j\left(F_{ij}\frac{1}{t_i}\frac{1}{\tau_j} - F_{ji}\frac{1}{t_j}\frac{1}{\tau_i}\right) = S_iS_j\left(F_{ij}A_{ij} - F_{ji}A_{ji}\right), \tag{11}$$

Where K is a scaling factor, and $A_{ij} = K/t_i\tau_j$. Equation (9) takes the same form as equation (3). From this it follows that the substitution parameters in FTT:Power, which are defined for each technology pair, are given by $b_{ij} = F_{ij}A_{ij} - F_{ji}A_{ji}$. Note that $b_{ij} = -b_{ji}$.

The shares equation, which determines the total rate of change of technology shares for technology i, is thus given by adding all the net flows:

$$\frac{dS_i}{dt} = \sum_j \frac{dS_{ij}}{dt} = \sum_j S_iS_j\left(F_{ij}A_{ij} - F_{ji}A_{ji}\right). \tag{12}$$

It follows from this equation that the total shares are always equal to one ($\sum_{ij} dS_{ij}/dt = 0$ and $\sum_i S_i = 1$).

Given initial shares (in 2013), and electricity demand (from 2013-2050), equation (12) is used in FTT:Power to compute global technology deployment in the power sector. Note that the shares equation is computed independently in each of the 59 regions that are represented by FTT. Based on the shares, it is relatively straight-forward to calculate the total installed capacity and electricity produced globally by each technology (as shown in figure 6.1) over time[76]. Total power sector emissions (as shown in figure 6.2) are then given by $E(t) = \sum_i \alpha_i G_i(t)$, where α_i are the emissions factors.

6.2.3 Policies in FTT:Power

With the exception of regulations and kick-start policies, all policies in FTT:Power are represented as changes to the LCOEs. In this sense, policies are seen to impact investor preferences and thus market dynamics. As already seen, carbon prices are represented by a variable in the LCOE (the carbon cost ($CC_i(t)$). Subsidies are modelled as reductions in the investment costs, TI_i, of technologies, and FiTs are modelled as effective subsidies, which cover the differences between the LCOE and the price of electricity plus a margin (Mercure et al., 2014). Lastly, regulations are modelled as a limit to the construction of new units of technologies and kick-start policies are modelled as a small increase in shares.

The impacts of different combinations of policies on technology deployment and emissions in FTT and E3ME-FTT have been explored in several papers (e.g. Mercure et al., 2014). This chapter instead asks how certain we can be that given policy sets, if implemented in the real world, will lead to the modelled results. An understanding of the uncertainties involved is key to assessing the predictive value of FTT:Power and evaluating the usefulness of simulation models compared to energy system optimisation models. It is also key to assessing the chance that model outputs might be wrong, which – as argued in Chapter 4 – is crucial in this area of research.

To properly assess the reliability of FTT predictions we would have to assess whether the shares equation represents an accurate description of technology dynamics and whether the impacts of policies are adequately captured. These are both crucial, but huge, questions. This chapter, instead, conducts a global sensitivity analysis of the FTT:Power parameters that define the core FTT equation presented in this section. Any interpretation of the results of the sensitivity analysis as representative of the uncertainty of policy impacts will implicitly assume not only that the distributions of parameter values used in the analysis represent the "true" uncertainty of parameter values, but also that the structure of FTT:Power offers an adequate representation of power sector dynamics. Given that this chapter only examines parametric uncertainty, and only do so for a subset of FTT:Power parameters, the results of

---

[76] See (Mercure, 2012) for the full set of equations.

the analysis can be viewed as a conservative estimate of the uncertainty of FTT:Power predictions (see discussion of parameter ranges in section 6.4.2). The method employed and the distributions of parameter values used in the analysis are presented in section 6.4. Before so doing, a review of previous sensitivity analyses of the FTT model and other IAMs is provided.

6.3 Past uncertainty assessments

The goals of the analysis conducted in this chapter are to i) provide a first conservative estimate of the uncertainty of FTT:Power predictions and, ii) to identify the parameters in the FTT core equation, the shares equation, that have the biggest influence on FTT predictions. The analysis represents a significant extension of earlier FTT sensitivity analyses because it employs a global instead of a one-factor-at-a-time approach (explained in the next section) and because it covers all the parameters that define the shares equation, several of which have not yet been included in sensitivity analyses of FTT.

The most comprehensive sensitivity analysis of FTT:Power thus far is found in Pablo Salas' PhD thesis (Salas, 2017). Salas considers the uncertainty of energy resource availability, learning rates, and power grid flexibility on power sector emissions. With the exception of one case study (of the uncertainty of the availability of hydroelectricity in Brazil), however, Salas considers only the extreme values of each parameter using a one-factor-at-a-time approach. In addition to Salas (2017), two papers presenting results from the full E3ME-FTT model include brief sensitivity analyses in the supplementary material (Mercure, Pollitt, Edwards, et al., 2018; Mercure, Pollitt, Viñuales, et al., 2018). In both cases, a one-factor-at-a-time approach based on only extreme values is used. Three FTT:Power parameters are varied by Mercure et al. (2018; 2018): renewables capital costs, renewables learning rates, and the investor discount rate. The sensitivity of only one FTT:Power output, the renewables technology share, is examined (i.e. the sensitivity of emissions or the deployment of other technologies to the three parameters is not reported) and found to be most sensitive to the investor discount rate. In all of the above examples, it is the sensitivity of the predicted impacts of '2C Policies' and 'No Policies' that are examined. No sensitivity analysis has so far been conducted of the predicted impacts of the '1.5C Policies'.

Uncertainty assessments of IAMs have been repeatedly called for in the literature (e.g. Kann & Weyant, 2000). The criticism of CBA IAMs, in many ways, boils down to the argument that the results are determined by parameters that are either unknown (climate damages) or highly value-laden (the discount rate). Most modelers working on CBA IAMs have recognized that uncertainty is a key challenge. As a consequence, there are numerous studies assessing the uncertainty of assumptions in CBA IAMs (Beck & Krueger, 2016).

Comprehensive uncertainty assessments are much rarer for large-scale IAMs than for CBA IAMs, in part due to their size and complexity. But several studies exist that assess uncertainties associated with large-scale IAMs focusing on select groups of parameters. Optimal pathways generated by IAMs have long been known to be highly sensitive to technology cost assumptions (Keepin & Wynne, 1984) and AR5 highlighted the importance of assumptions regarding both the availability and costs of future technologies for mitigation costs (IPCC, 2014a). The EMF 27 study (Kriegler et al., 2014) and the AMPERE Project (Riahi et al., 2015a) were both specifically dedicated towards assessing the impacts of technological uncertainty on the cost of mitigation. Overall, the IAM community has studied the influence of technological and socio-economic assumptions on low-carbon transformation pathways and mitigation costs extensively (e.g. Krey and Riahi (2009), van Vliet et al. (2009), Edenhofer et al. (2010), McJeon et al. (2010), and Bosetti et al. (2015)).

The outputs that are examined in most IAM uncertainty assessments, however, differ from the outputs analysed in FTT:Power in this chapter. This has to do with the "mode" in which IAMs are run, as discussed in Chapter 5. Because the majority of IAMs identify cost-effective emissions pathways that meet predetermined climate targets – typically utilizing techniques of constrained optimisation where emissions is one of the key binding constraints – emissions normally remain constant during sensitivity analyses. IAMs might, for example, be run in an climate constrained scenario with and without the availability of key technologies and with a range of parameter values in order to assess the increase in mitigation costs or carbon prices (Bosetti et al., 2015; Kriegler et al., 2014). The EMF 27 study for example, which "investigated the importance of individual mitigation options such as energy intensity improvements, carbon capture and storage (CCS), nuclear power, solar and wind power and bioenergy for climate mitigation" (Kriegler et al., 2014, p. 353), did so by examining the impact of varying these options on the cost of mitigation. Bosetti et al. (2015) similarly investigated the sensitivity of mitigation costs to energy technology costs (informed by expert elicitations) in climate constrained scenarios in WITCH, GCAM, and MARKAL. The sensitivity of *emissions* to assumptions was only investigated in the baseline scenario, which is the only scenario in which emissions are unconstrained and thus allowed to vary. Overall, the output whose uncertainty is most often analysed in uncertainty assessments of IAMs is the cost of mitigation (e.g. Rogelj et al. (2013), Iyer, et al. (2014), Barron and McJeon (2015), Olaleye and Baker (2015), McJeon et al. (2010), Lemoine and McJeon (2013), Lehtveer and Hedenus (2015)).

Thus, the sensitivity of emissions in non-baseline (mitigation) scenarios to assumptions is less often explored in IAM uncertainty assessments. As discussed in Chapter 5, when IAMs are run in "optimisation mode" (or, more or less equivalently, in "backcasting" mode) emissions are inputs and variables such as the carbon price and the cost of mitigation are outputs. The exception is when these IAMs are run without the emissions constraint, to generate baselines. In this case, the effect of varying input assumptions on the model's technology choice will also have an impact on emissions (e.g. Bosetti et al. (2015)).

Examples of analyses in which the sensitivity of emissions to input assumptions are investigated can, however, be found if we look to sensitivity analyses of simulation models such as IMAGE (Campolongo & Braddock, 1999; Van der Sluijs et al., 2002). The sensitivity analysis conducted in this thesis is much closer to these than it is to the preceding examples because FTT:Power is always run in "simulation mode", meaning that policies (the input variables) are given while technology deployment and emissions (the output variables) are allowed to vary not only in the baseline, but in all scenarios.

Lastly, it should be noted, despite efforts to examine uncertainties in IAMs, fewer studies have, according to Usher (2016), examined uncertainties in ESOMs (which in this thesis are treated as a sub-category of IAMs, see Chapter 3). This is despite a long-standing strong awareness of the importance of structural and parametric uncertainty in ESOMs (Keepin & Wynne, 1984). Yue et al. (2018) express concerns that uncertainties in model structures and input parameters in ESOMs are underplayed or ignored. Based on a comprehensive review of 2100 studies, Yue et al. identify little over 100 studies that use deterministic scenarios to explore uncertainty and only 34 studies that apply formal uncertainty techniques, nine of which use MCA. This is despite MCA being the most well-known method of uncertainty analysis (Anadon et al., 2017) (other formal techniques identified by Yue et al. are stochastic programming, robust optimization, and modelling to generate alternatives). Thus, uncertainty assessments of energy system models that go beyond scenario analysis are still relatively uncommon. In a recent review of the IAM literature, Gambhir et al. (2019) also write that it is (still) not common practice to delve into the assumptions that drive IAM results and that more frequent used of sensitivity analyses would be beneficial for IAM research. The analysis conducted in this chapter is partly a response to this.

6.4 Method

6.4.1 Global sensitivity analysis

Sensitivity analyses are often conducted using a one-factor-at-a-time approach (Morris, 1991), that is, by varying one parameter at a time and recording outputs while keeping all other parameters constant. Often, when using a one-factor-at-a-time approach, only the extreme values of input parameters are considered (which, as shown in the previous section, has been the case for almost all of the sensitivity analyses conducted of FTT thus far). A global sensitivity analysis (GSA), by contrast, is based on varying parameters *simultaneously* over their entire ranges (Johnson et al., 2006; Saltelli et al., 2006)[77]. A GSA is preferable for two reasons. First, it captures potential interaction effects between parameters. Second, by capturing the entire ranges of input values, it ensures that the entire ranges of output values (corresponding to the input values) are captured, including when the extreme values are found on the interior of the input ranges. Both are important when analysing the sensitivity of complex models that exhibit strong non-linearities such as FTT:Power (Saltelli et al., 2006).

The GSA in this chapter is conducted using a Monte Carlo analysis (MCA). In an MCA, distributions of inputs are sampled at random to compute the corresponding distributions of outputs. An MCA thus provides a picture of the variability of outputs conditional on the variability of inputs. An additional benefit of MCA, compared to other sensitivity analysis methods, is that the accuracy of the analysis depends on the sample size, but not on the number of input parameters (Morgan & Henrion, 1990). MCA is the most well-known method for uncertainty analysis (Anadon et al., 2017) and it is well suited for GSA.

In conventional MCA, inputs are sampled completely at random from their distributions. For each new run, parameters are drawn independently from previous runs. This procedure can, however, lead to clustering of parameter values, which means that other regions of the input parameter space are less well represented (Loucks et al., 2005). Because the aim of the analysis conducted in this chapter is to better reflect the variability of outputs contingent on the variability of inputs, such clustering is best avoided. A stratified sampling method is useful in this context because it ensures an even coverage of the input parameter space (Loucks et al., 2005).

Latin Hypercube Sampling (LHS) represents a commonly used stratified sampling method (Iman & Conover, 1982). LHS divides each input distribution into section of equal probability before drawing one value at random from each section. Values from each section for each parameter are randomly assigned to values from sections from the other parameters. Because LHS reflects the mean, variance, and other aspects of the input distributions more accurately (for the same sample size) than completely random sampling, it represents a very efficient sampling method (Morgan & Henrion, 1990). This means that the distribution of outputs also reflects the distribution of inputs more accurately (for the same sample size). LHS has been used in several uncertainty analyses of the CBA IAMs DICE, RICE, and PAGE (Butler et al., 2014; Stanton et al., 2008). LHS is used more recently by Chan and Anadon

[77] A 'global' sensitivity analysis is used in this thesis in line with Johnson et al. (2006) and Saltelli et al. (2006) to denote the perturbation of multiple model inputs simultaneously (as opposed to individually) and the evaluation of the effects of doing so on model outputs. 'Global' does not mean that *all* the parameters in a model are included in the sensitivity analysis.

(2016) to propagate uncertainty through MARKAL to estimate the benefits of R&D portfolios, and Anadon et al. (2017) review numerous efforts to incorporate approaches, including LHS, to facilitate public R&D decision-making under uncertainty.

6.4.2 Selected parameters and their distributions

The rate of technology deployment (and thus emissions reductions, since electricity demand is exogenous) in FTT:Power is determined by the shares equation (equation 12). It follows from this that the rate of growth of a given technology depends on the market shares of all the other technologies and on the substitution parameters between the given technology and all the other technologies. The dependence of the rate of growth on the existing market shares of all technologies is what gives rise to the characteristic shape of diffusion (the S-shape). A technology with a very small market share will grow slowly because the rate at which it can replace other technologies is proportional to its own share. However, if the market share of a technology becomes very large, approaching unity, all the other shares necessarily approach zero, which again limits the rate of growth of the technology with a large share. Thus, the form of the shares equation determines the shape of technology deployment. The rates of deployment, given shares, however, are determined by the substitution parameters, $b_{ij} = F_{ij}A_{ij} - F_{ji}A_{ji}$. The global sensitivity analysis conducted in this chapter targets the parameters that define the substitution parameters, which are similar for all FTT models. Due to time constraints of this doctoral research, remaining FTT:Power parameters are left for future sensitivity analyses. Appendix E lists parameters in FTT:Power that are not included in the sensitivity analysis conducted in this chapter. The fact that not all parameters are included means that the results presented do not capture the full variability of FTT:Power outputs and thus likely represent (see discussion of parameter ranges below) a conservative estimate of the uncertainty of FTT:Power predictions.

This chapter includes five groups of parameters: three related to A_{ij} and two to F_{ij}. The parameters that define A_{ij} are the characteristic time constants: the lifetimes, τ_j, build times, t_i and overall scaling factor, K. None of these parameters have been investigated in sensitivity analyses of FTT:Power thus far. The parameters that define F_{ij} include the parameters that define the LCOEs and the learning rates. (In addition to these, F_{ij} also includes policy variables. Because the point of the sensitivity analysis is to get a better understanding of the uncertainty of the impacts of given policies, however, these variables are held constant.) The cost parameters that are already treated as distributions in FTT:Power are not included (that is, the investment costs, operation and maintenance costs, and fuel costs). This leaves us with two parameters to investigate: the investor discount rate, r, which determines how costs are aggregated over time, and the learning rates, LR_i, which determine how costs decrease with total installed capacity.

Because all of the parameters except for the scaling factor are technology specific, and there are 24 technologies in FTT:Power, this gives rise to 97 different parameters (96 technology-specific and one global). Table 6.1 lists the parameters and their impact on the rate of technology deployment.

Table 6.1 Parameters included in the global sensitivity analysis

Parameters		Impact on the rate of technology deployment	Global or technology specific variation
τ_j	Lifetimes	Longer lifetime of a technology makes it harder for other technologies to replace it	Technology specific
t_i	Build times	Longer build time of a technology makes it harder for it to replace other technologies	Technology specific
K	Overall scaling factor	Higher values of K will speed up technology diffusion overall	Global
r_i	Investor discount rates	A higher discount rate for a technology makes it cheaper	Technology specific
LR_i	Learning rates	A higher learning rate for a technology implies faster cost reductions over time	Technology specific

For the purpose of the GSA conducted in this chapter, the parameter values are assumed to be independently and uniformly distributed around ±50% of their default values. All the parameter values, including the default and ±50% ranges, are shown in Table 6.1.

Table 6.2 Best guess parameter values and ±50% ranges (in parenthesis) used in MCA.

Technology	Lifetimes τ_j [years]	Build times t_i [years]	Investor discount rates r_i	Learning rates LR_i [%]	Overall scaling factor K
Nuclear	60 (30, 90)	7 (3.5, 10.5)	0.1 (0.05, 0.2)	5.8 (2.9, 8.7)	
Oil	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	1.0 (0.5, 1.5)	
Coal	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	3.0 (1.5, 4.5)	
Coal + CCS	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
IGCC	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	3.0 (1.5, 4.5)	
IGCC + CCS	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
CCGT	30 (15, 45)	2 (1, 3)	0.1 (0.05, 0.2)	4.0 (2.0, 6.0)	
CCGT + CCS	30 (15, 45)	2 (1, 3)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
Solid Biomass	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
S Biomass CCS	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	7.0 (3.5, 10.5)	
BIGCC	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
BIGCC + CCS	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	7.0 (3.5, 10.5)	4 (2, 8)
Biogas	30 (15, 45)	2 (1, 3)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
Biogas + CCS	30 (15, 45)	2 (1, 3)	0.1 (0.05, 0.2)	7.0 (3.5, 10.5)	
Tidal	80 (40, 120)	7 (3.5, 10.5)	0.1 (0.05, 0.2)	1.4 (0.7, 2.1)	
Large Hydro	80 (40, 120)	7 (3.5, 10.5)	0.1 (0.05, 0.2)	1.4 (0.7, 2.1)	
Onshore	25 (12.5, 37.5)	1 (0.5, 2)	0.1 (0.05, 0.2)	7.0 (3.5, 10.5)	
Offshore	25 (12.5, 37.5)	1 (0.5, 2)	0.1 (0.05, 0.2)	9.0 (4.5, 13.5)	
Solar PV	25 (12.5, 37.5)	1 (0.5, 2)	0.1 (0.05, 0.2)	17.0 (8.5, 25.5)	
CSP	25 (12.5, 37.5)	1 (0.5, 2)	0.1 (0.05, 0.2)	10.0 (5.0, 15.0)	
Geothermal	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	5.0 (2.5, 7.5)	
Wave	20 (10, 30)	1 (0.5, 2)	0.1 (0.05, 0.2)	14.0 (7.0, 21.0)	
Fuel Cells	20 (10, 30)	2 (1, 3)	0.1 (0.05, 0.2)	15.0 (7.5, 22.5)	
CHP	40 (20, 60)	4 (2, 6)	0.1 (0.05, 0.2)	3.0 (1.5, 4.5)	

The ±50% range is commonly used in sensitivity analyses of IAMs. In a comprehensive review of 2100 ESOM studies, Yue et al. (2018) identify only nine studies that use MCA to assess uncertainty. In seven of these, input parameter ranges are defined by multiplying default values by an arbitrary percentage (both Lehtveer and Henedus (2015) and Pye et al. (2015) use ±50% ranges for all or some parameters in their sensitivity analyses) and the distributions tend to be simple (uniform or triangular). (Only two of the studies (Bosetti et al., 2015; Fragkos et al., 2015) offer more comprehensive assessments of input parameter ranges and distributions.) Van der Sluijs et al. (2002) also use the ±50% range in their global sensitivity analysis of the TIMER model. Uniform distributions are often used when little is known about a distribution; The more "complicated" the shape of a distribution is, the more information is needed to specify it[78].

A more comprehensive assessment of FTT:Power input parameter ranges and distributions is beyond the scope of this thesis. For the purpose of identifying the parameters in the FTT shares equation that have the largest impact on FTT:Power predictions, however, the parameter distributions (i.e. the ranges and shapes) do not matter as long as the sensitivities are measured in relative terms, which they are in this thesis (using Pearson's correlation coefficient). The recommendation in this chapter is that future uncertainty analyses of FTT:Power consider the uncertainty of the parameters that – based on the sensitivity analysis in this chapter – turn out to have a larger impact on FTT:Power predictions (the parameter with the largest impact on predictions, we will see, is the overall scaling factor). For the purpose of providing a first *conservative* estimate of the uncertainty of FTT:Power predictions, it is enough to show that the ranges obtained by varying the best guess parameter values by ±50% are not unreasonably large[79]. Appendix D compares the ±50% ranges for learning rates, lifetimes, and investor discount rates with ranges found in the literature, reflecting either empirical estimates or ranges used by other modelers. Even though the information in Appendix D is not based on a comprehensive review, the findings indicate that the ±50% ranges are within the bounds of what is believed to be possible (a more comprehensive review can only increase the ranges found in the literature). The scaling factor is not reviewed because there is no literature on it (it is a parameter that is specific to FTT, which is not even discussed in FTT publications) and no information was gathered for the build times, whose unclear specification is discussed in the next section.

A caveat of the sensitivity analysis conducted in this chapter is that correlations between parameter values are not taken into account. The independence assumption may in general both underestimate and overestimate the uncertainty in outputs (Parkinson & Young, 1998; Shackley et al., 1998). The fact that the directions of the impacts of parameters on the rates of technology deployment and thus emissions are known in FTT:Power (Table 6.1), however, enables us to make an educated guess regarding the effects of the independence assumptions. First, it is not implausible that parameters such as e.g. the hurdle rates (investor discount rates) are correlated for renewable technologies. If this is the case, it means that situations in which some renewable hurdle rates are high and some are low are less likely than situations in which all renewable hurdle rates are either high or low. This, it seems, will increase

---

[78] The uniform distribution corresponds to the maximum entropy probability distribution if we know only the minimum and maximum values.

[79] The fact that only a sub-set of FTT:Power parameters are included and that structural uncertainty is ignored also contributes to the conservatism of the estimates in this chapter. It is possible that even with relatively large parameter ranges, the results of the analysis will still represent a conservative estimate of the uncertainty.

the frequency of high and low emissions outcomes relative to median emissions outcomes, i.e. the variance of the distribution of emissions, which may be interpreted as an increase in the uncertainty of emissions. This is because when all renewable hurdle rates move together this has a stronger impact on emissions (in either direction) than when renewable hurdle rates move in different directions. In general, for FTT:Power, if hurdle rates, build times, lifetimes, or learning rates are correlated for either low-carbon technologies as a group or for fossil fuel technologies as a group, this seems to increase the frequency of high and low emissions outcomes relative to median emissions outcomes. Thus, it is not unlikely that in FTT:Power, the independence assumption leads to an underestimate of the uncertainty of emissions (if, however, parameters for low-carbon technologies are correlated with parameters for fossil fuel technologies, the opposite might be the case). Although future uncertainty analyses of FTT:Power would benefit from considering potential correlations, the assumption does not stop us from using the sensitivity analysis to identify the individual parameters that influence emissions and technology deployment in FTT:Power the most. Additionally, given little reason to believe that the independence assumption leads to a vast exaggeration of the uncertainty of technology deployment and emissions, the results of the GSA can still be interpreted as a first conservative estimate of the uncertainty of FTT:Power predictions.

It is nonetheless important to highlight that the analysis conducted in this chapter represents only the first step towards a more comprehensive uncertainty assessment of FTT:Power and that the uniform and independent distributions represent only first order approximations of the uncertainties associated with selected FTT:Power parameters. The advantage of starting with a relatively simple global sensitivity analysis is that future efforts can be directed towards those parameters that turn out to have a large impact on model results. When it comes to potential correlations between parameters, it might be worthwhile trying out possible values and see whether this affects results. Based on this, efforts can again be directed towards those correlations that matter for FTT:Power predictions.

6.4.3 Uncertainty of selected parameters

An uncertainty analysis goes beyond a sensitivity analysis because it does not only compute the effects of varying input parameters on model outputs but interprets the variability as representative of uncertainty. When the input distributions used in the MCA are seen to reflect the uncertainty of input parameter values, the output distributions can be interpreted as the uncertainty of outputs.

The parameter distributions used in this chapter, however, do not, as already discussed, reflect the 'true' uncertainty of the input parameters. In order to determine what the 'true' uncertainty of the input parameters are, more research needs to be done. Uncertainty, however, will always remain subjective in the sense that some people will know more than others and more research might reduce it (Morgan

& Henrion, 1990). In that respect, the uniform distributions employed in this chapter can be interpreted as the "maximum ignorance" estimates of the uncertainty of the selected parameter values in FTT:Power. The brief comparison of parameter distributions used in this chapter and the ranges found in the literature (shown in Appendix D and discussed in the previous section) also suggests that the ±50% ranges are within the bounds of what is thought to be possible.

In any case, the GSA computes the variability of FTT:Power predictions contingent on specified input distributions. The results of the analysis can be used to prioritise future uncertainty analyses, in which more attention can be paid to capturing the 'true' uncertainty of the parameters that have a large impact on results. A qualitative discussion of the uncertainty of the selected parameters is included here.

The uncertainty of the build times (t_i) and lifetimes (τ_j) in FTT:Power has two main sources. First, the data (obtained from (IEA & NEA, 2015) is not perfect. Both build times and lifetimes vary for different units of the same type of technology, between regions, and over time. In addition to that, lifetimes of power plants can be actively extended and build times might be actively shortened (more on the ability to control parameter values in section 6.6). Second, there is uncertainty regarding the appropriate specification of these parameters in FTT:Power. This uncertainty arises partly from the novelty of the FTT framework, which means that different specifications have not yet been tested. In particular, it is unclear whether the construction times of different technologies reported by the IEA are the right values for the t_i parameters in FTT:Power. Because t_i represent the "birth rates" of technologies (Mercure, 2015), an equally plausible measure might be the time from the inception of a new power plant until the time it starts operating. This would, however, be different from the construction times of different technologies (which are currently used). Changing this measure would likely increase the build times of more unpopular technologies (such as nuclear). We might also expect that build times measured in this way are more likely to change with experience over time. In summary, thus, the uncertainty of build times includes both the uncertainty of the IEA data and the uncertainty regarding the correct specifications according to the underlying theory. While more and better data would reduce the first type of uncertainty, this would not reduce the second type of uncertainty.

Of all the selected parameters the scaling factor (K) is the one that is characterised by the most fundamental uncertainty. The value of this parameter is not discussed in FTT publications. It is based, however, on a judgment of how long it takes to achieve a full turnover of technologies. More specifically, the value of the scaling factor is chosen so as to match Arnolf Grübler's timescale of a full turnover every 100 years[80].

[80] Personal communication with Jean-Francois Mercure

In the limited sensitivity analyses that have been conducted of FTT:Power so far the investor discount rate (r) had the largest impact on results (Mercure, Pollitt, Edwards, et al., 2018; Mercure, Pollitt, Viñuales, et al., 2018). The investor discount rate, also known as the hurdle rate, represents the minimum rate of return that an investor needs to earn in order to make an investment. In FTT:Power, the investor discount rate is set to 10% for all technologies in all regions. Recent literature, however, has shown that investor discount rates vary based on the technology in question (Egli et al., 2018) and from region to region (Ondraczek et al., 2015) due to differences in risk perceptions and the cost of capital. At the same time, it is extremely difficult to obtain data from investors on the discount rates they employ (due in part to commercial reasons). While previous sensitivity analyses of FTT:Power only investigated the impacts of changing r by the same factor for all technologies, the analysis in this chapter captures technology-specific variations of the investor discount rates, r_i.

When it comes to the learning rates, LR_i, it has long been known that the assumptions of technology costs and dynamics have a large impact on mitigation costs (Löschel, 2002). The inclusion of endogenous learning will, everything else being the same, lead to significant cost reductions. AR5 highlights the importance of technology cost assumptions for mitigation costs (IPCC, 2014a) and the analysis of these aspects has since been at the centre of a growing literature (Bosetti et al., 2015). FTT:Power is unique in this regard because it allows us to investigate the impacts of learning rates on technology deployment, and thus emissions, as opposed to the cost of meeting a given target.

6.5 Results

6.5.1 Technology deployment

The results of the MCA (for n=200 runs) for the deployment of key FTT:Power technologies given 1.5C policies and 2C policies are shown in figures 6.5 and 6.6 respectively. The technologies shown include the five technologies that saw the largest changes in the best guess predictions compared to the baseline: Coal, Gas, Solar PV, Wind, and NETs (which in FTT:Power consist of three kinds of bioenergy with CCS (BECCS)). The figures show a wide range of trajectories surrounding the best guess predictions for each of the five technologies. If we interpret the spread of trajectories as the uncertainty of technology deployment (contingent on the assumed uncertainty of input parameters), this indicates a very large uncertainty with regards to the impacts of policies on technology deployment.

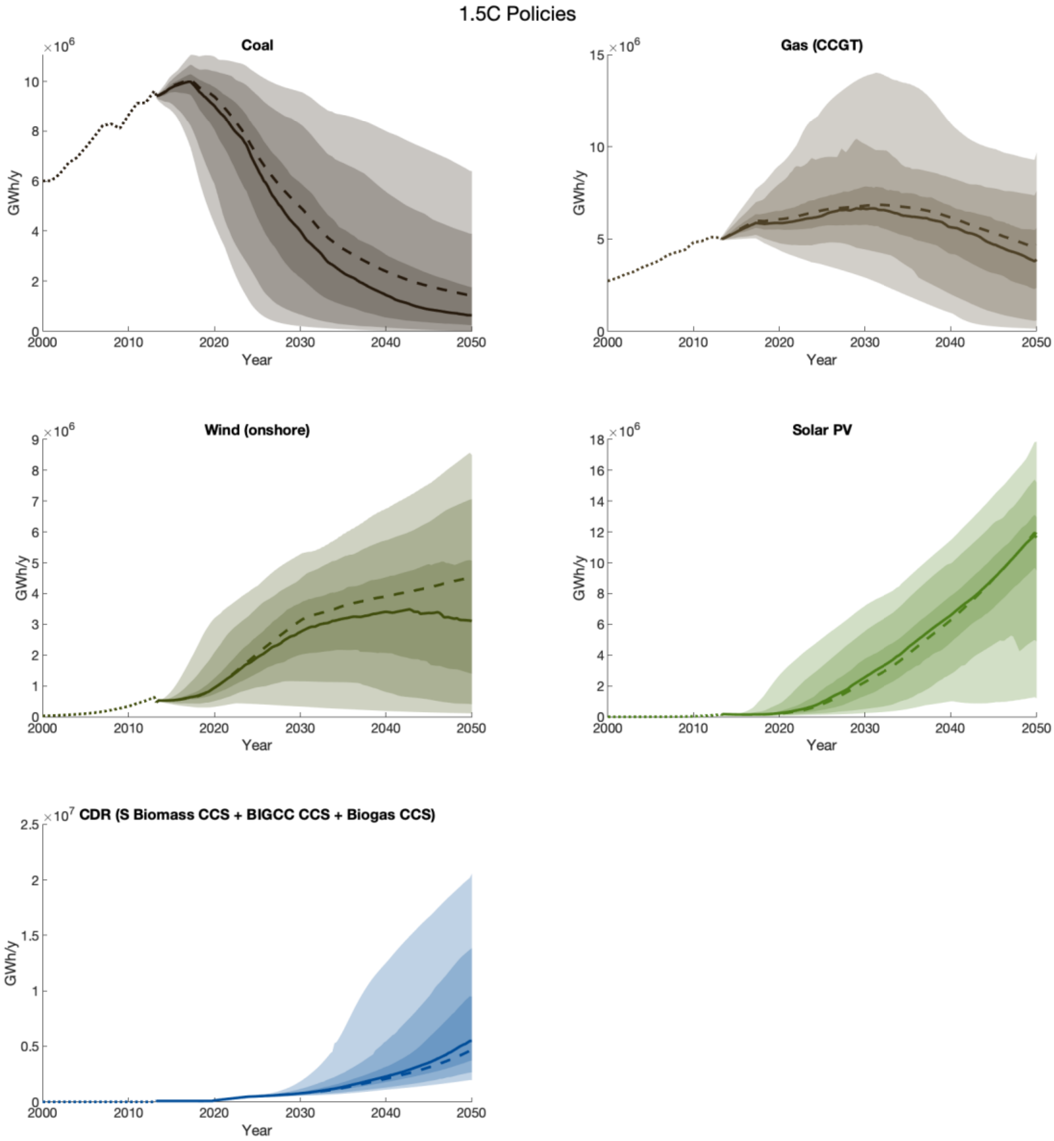


Figure 6.5 Ranges of MCA runs (n=200) for electricity generation for coal, gas, wind, solar PV, and CDR (NETs) in the 1.5C scenario. The lightest shade shows the full range, and the darker shades show the 5-95% range and the 25-75% range. The solid line shows the 50% and the dashed lines show the best guess predictions.

Figures 6.5 and 6.6 also show that the uncertainty of technology deployment caused by policies increases over time. This is in accordance with expectations: the further into the future we look, the harder it is to predict the impacts of policies. Figures 6.7 and 6.8 show more clearly how the distributions of technology deployment in the MCA runs evolve over time. While figure 6.6 also show that distributions generally widen over time, it also shows that the deployment of coal in the MCA runs is more spread out in 2030 than in 2050. This can be explained by the fact that the 2C policy set includes regulations that are used to phase out and cap coal in some regions. Because regulations are modelled as hard limits on shares, they necessarily reduce the uncertainty of technology deployment (of the regulated technologies).

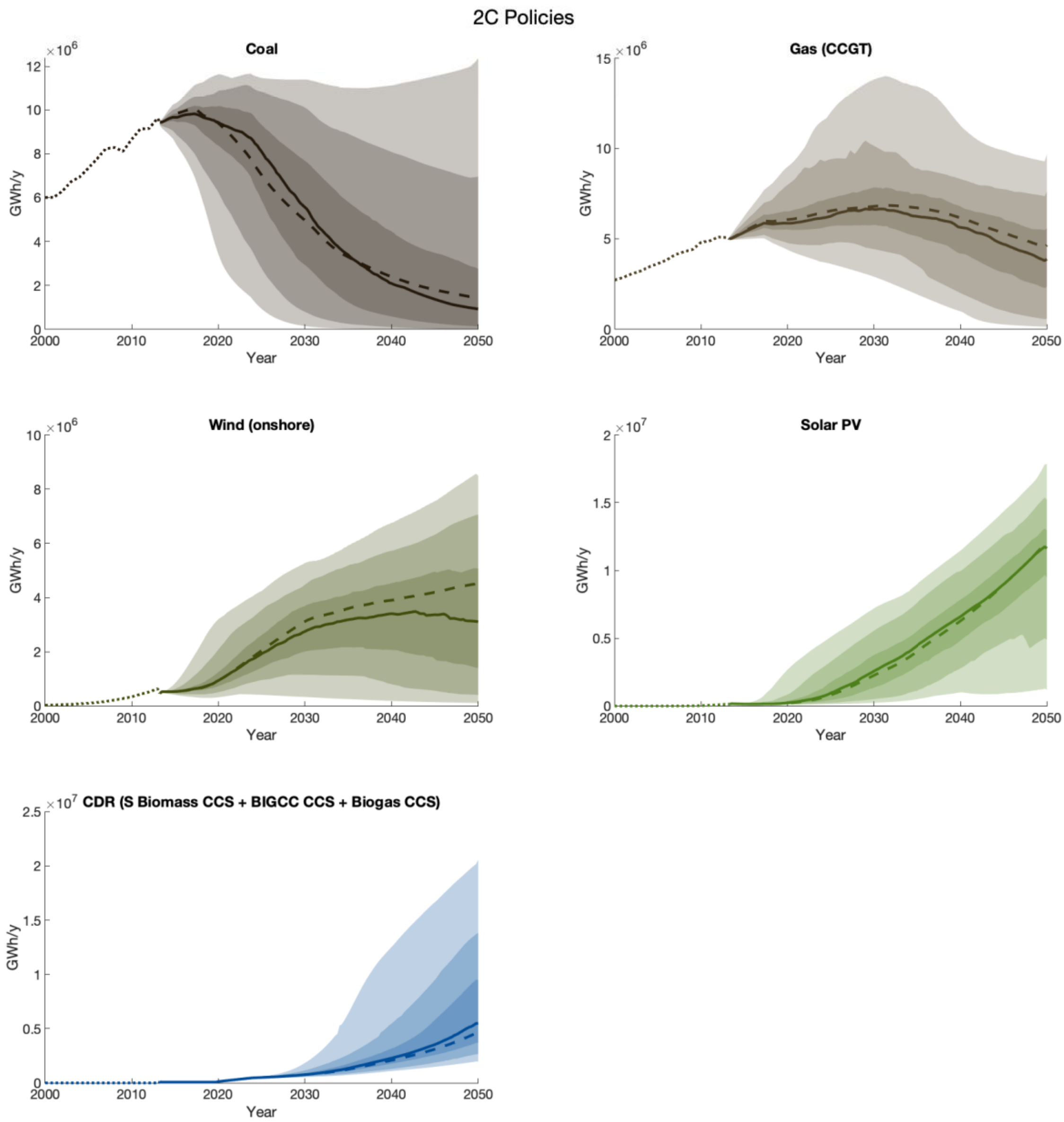


Figure 6.6 Ranges of MCA runs (n=200) for electricity generation for coal, gas, wind, solar PV, and CDR (NETs) in the 2C scenario. The lightest shade shows the full range, and the darker shades show the 5-95% range and the 25-75% range. The solid line shows the 50% and the dashed lines show the best guess predictions.

Figures 6.7 and 6.8 also show a large spread in the level of technology deployment for most technologies already in 2030. Despite narrower ranges in absolute terms for renewable technologies, their coefficients of variation (i.e. the standard deviations divided by the means) are comparable to other technologies. For coal, the level of deployment is already significantly spread out in 2020, which is only 7 years after the start date of FTT:Power simulations.

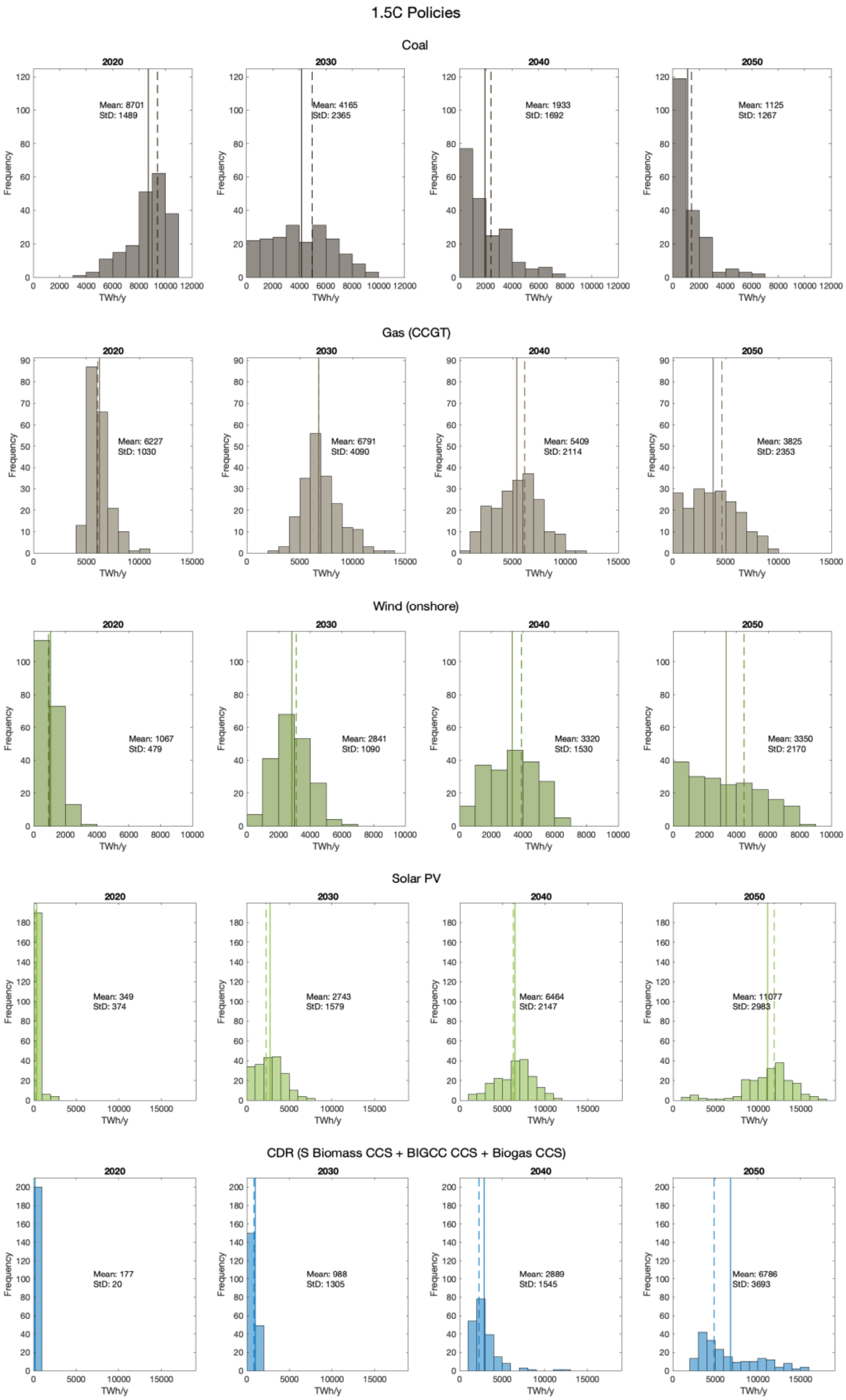


Figure 6.7 MCA runs (n=200) for electricity generation at different points in time for the 1.5C scenario. The dashed lines show the best guess prediction and the solid lines show the median values of the Monte Carlo runs.

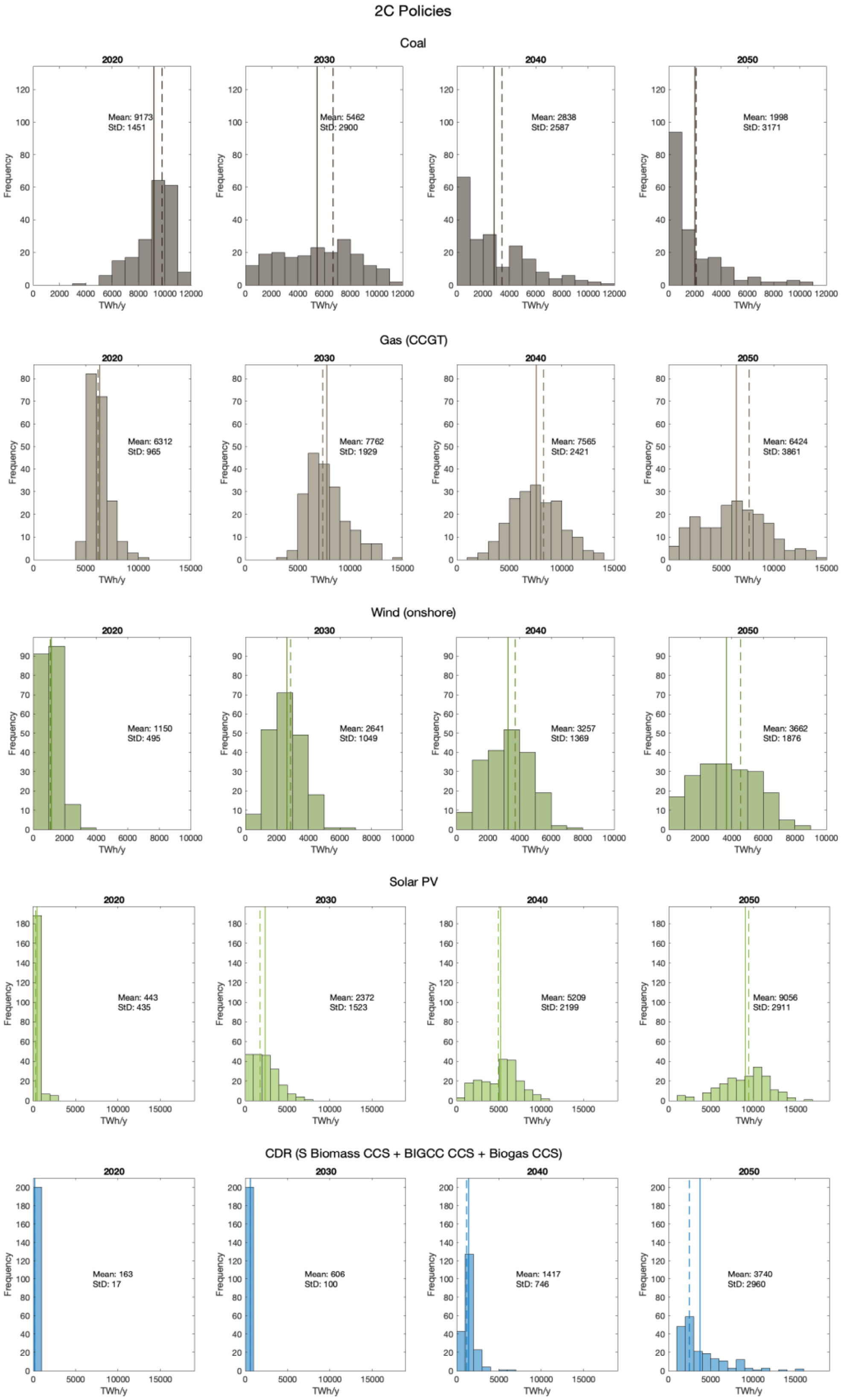


Figure 6.8 MCA runs (n=200) for electricity generation at different points in time for the 2C scenario. The dashed lines show the best guess prediction and the solid lines show the median values of the Monte Carlo runs.

Figure 6.9 shows the same results interpreted in probabilistic terms for the 1.5C runs. The figure shows the probabilities of generating less than x TWh/y in the case of coal and gas, and more than x TWh/y in the case of wind, solar, and NETs. The probabilities are defined as the fraction of Monte Carlo runs below and above x TWh/y.

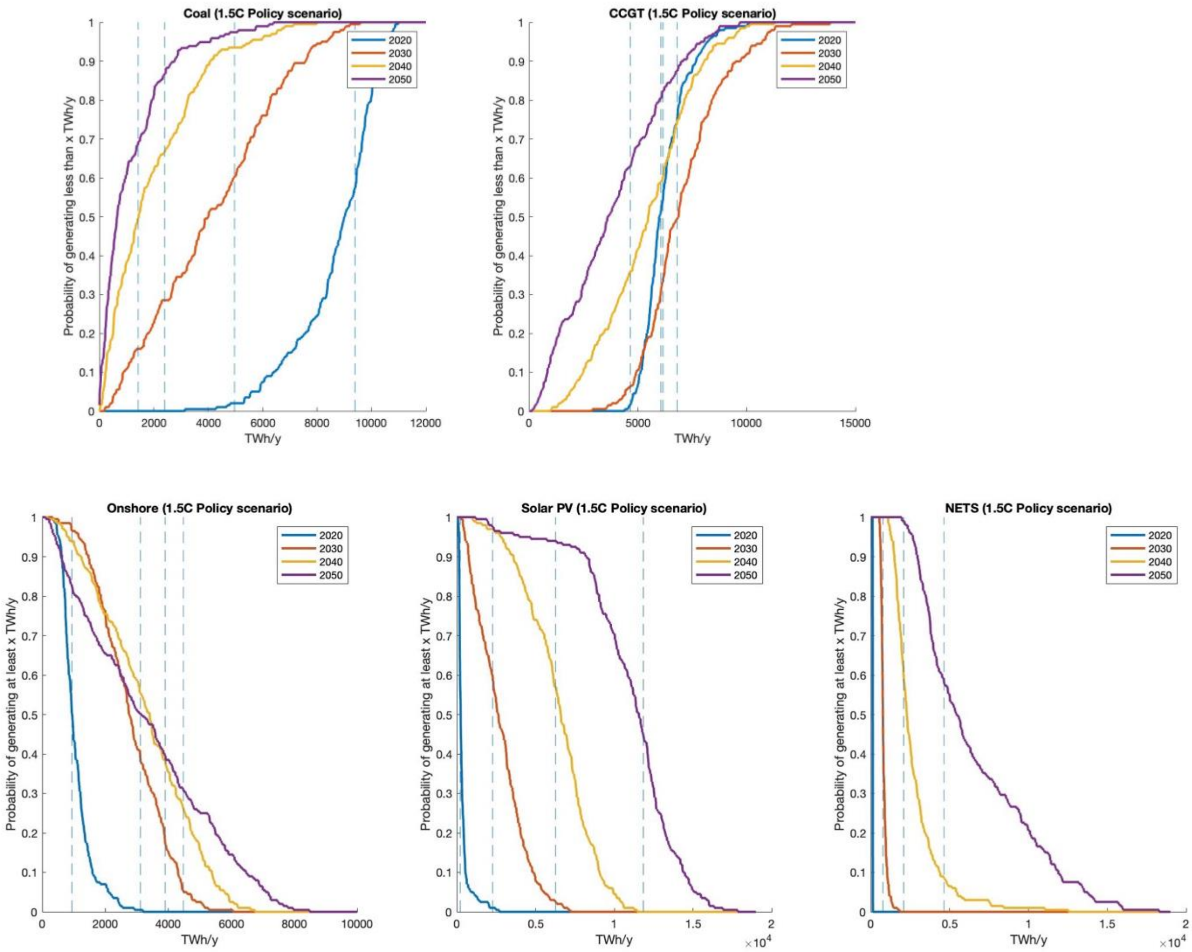


Figure 6.9 Cumulative probability distributions for generating less than and more than a given level of electricity (TWh/y) calculated as fractions of MCA runs in the 1.5C scenario. Dashed lines indicate best guess levels of electricity generation in the corresponding years.

Since FTT:Power is used (as part of E3ME-FTT) to inform energy and climate policy, a key question is how well the model can predict the impacts of policies on technology deployment in the power sector. While we might not expect to be able to predict policy impacts many decades into the future, we might expect to be able to predict the impacts of policies in the near term. The parametric uncertainty of model outputs represents only a part of the total uncertainty. Even then, the results of the MCA analysis conducted here indicate that the uncertainty of policy impacts on technology deployment is high. The next section presents the results of the MCA for emissions.

6.5.2 Uncertainty of power sector emissions

Figure 6.2 showed the best guess predictions from FTT:Power for the impacts of policies designed to meet the 1.5˚C and the 2˚C targets on power sector emissions. Figure 6.10 shows the spread of emissions trajectories obtained from the MCA for the two policy scenarios and the baseline. This figure provides a very different image of policy impacts compared to figure 6.2. If we take the uncertainties of core parameters into account, a wide range of emissions trajectories result. The first finding, thus, is that predictions of power sector emissions from FTT:Power are sensitive to selected core FTT parameters. At the same time, however, taking parametric uncertainty into account does not imply that "anything is possible". It is still clear from figure 6.10 that few of the No Policies runs reach levels of emissions comparable to the best guess predictions for the 2C and 1.5C policy sets. The 'worst cases', i.e. the runs with the highest emissions, also involve much lower emissions when policies are implemented. The fact that FTT:Power is based on diffusion theory means that new technologies can increase their markets shares even when no policies are in place, especially if learning leads to large cost reductions. It is therefore possible in principle that some of the baseline runs could also lead to a fast diffusion of low-carbon technologies. The MCA, however, indicates that the effects of policies on emissions in FTT:Power is robust with respect to the uncertainty of core parameters. This is in line with the finding in AR5 that policies are necessary to reduce emissions (IPCC, 2014a). Thus, the second finding is, if we believe that FTT:Power sufficiently captures the dynamics of technology diffusion in the electricity sector, parameters would have to take on very extreme values in order for emissions in the baseline to reach levels near those that are required to meet the Paris target.

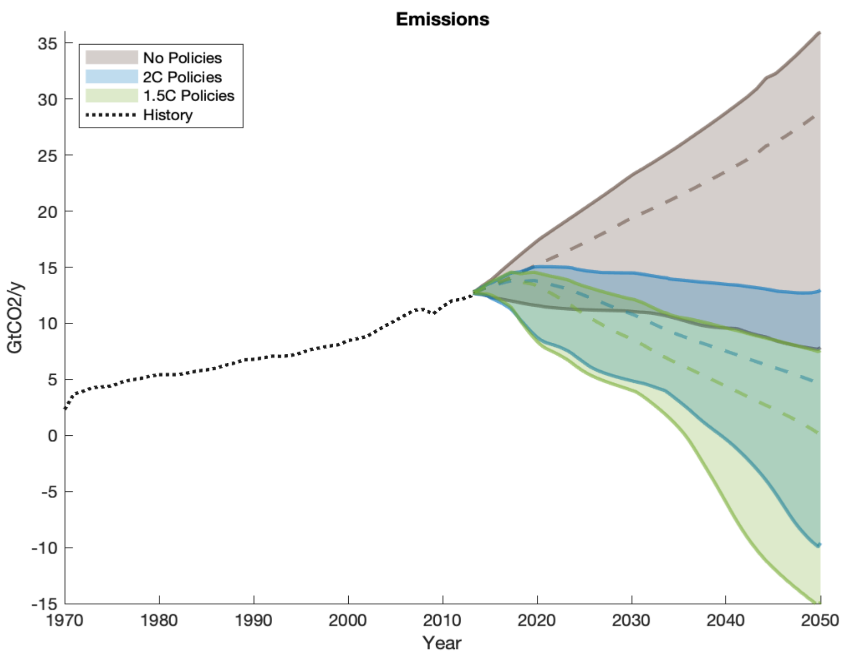


Figure 6.10 Range of emissions in MCA runs for the No Policies policy set (brown), 2C policy set (blue), and 1.5C policy set (green). The ranges indicate the maximum and minimum values. The dashed lines show the best guess predictions. (For each policy scenario n=200.)

Figure 6.11, which shows the spread of cumulative emissions in the MCA runs for each policy scenario, illustrates the second finding even more clearly. Figure 6.11 also shows that, while almost all baseline runs result in cumulative emissions that are higher than all the policy runs, there is a significant overlap in cumulative emissions between 2C policy runs and 1.5C policy runs. This leads to the third finding: the certainty with which we can distinguish the impacts on emissions of policies designed to meet 1.5˚C and 2˚C using FTT:Power is limited.

Table 6.3 Properties of MCA runs

Policy Set	**Default values ($GtCO_2$)**	**Mean ($GtCO_2$)**	**Standard Deviation ($GtCO_2$)**	**Runs below best guess (#)**	**Runs above best guess (#)**
No Policies	749	703	89	125	75
2C Policies	372	337	75	136	64
1.5C Policies	292	254	74	142	58

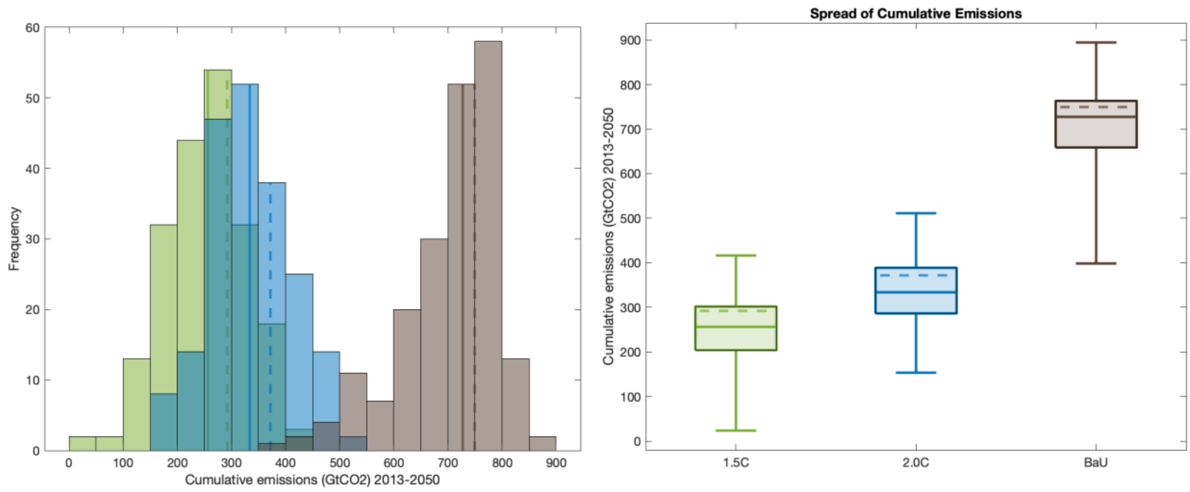


Figure 6.11 Histograms (left) and Box plots (right) showing cumulative emissions in the MCA runs for the No Policies (brown), the 2C (blue), and the 1.5C (green) Policy scenarios (n=200). Dashed lines show best guess predictions of cumulative emissions.

The overlap in cumulative emissions in the MCA runs for the 1.5C and 2C policy sets indicates that policies aimed at reducing emissions to levels that are compatible with 1.5˚C might lead to emissions that are closer to a 2˚C target. Of course, the level of global warming depends on the total emissions from all sectors, not just the power sector. (Furthermore, even if we knew the total emissions, there are significant uncertainties involved in predicting the impacts of emissions on temperature increases.) At the same time, power sector emissions represent the single largest source of global emissions (about 32% in 2018). Given the huge environmental impacts of a difference in global warming between 1.5˚C and 2˚C, even a fraction of this is likely to have large detrimental environmental consequences. The uncertainty of power sector emissions, therefore, should not be taken lightly when using models to inform policies and long-term strategies. As argued in Chapter 4, even a small chance of being wrong implies great risks.

In order to summarise the uncertainty associated with predicting policy impacts, two simple indicators are constructed. These indicators summarise the extent to which emissions might deviate from best guess predictions if we take into account the uncertainty of core parameters. The first indicator, O, represents a simple measure of overlap, which tells us the extent to which the impacts of different policy sets are distinguishable. This is given by the fraction of runs in the shaded area of the histogram in Figure 6.11. The second indicator, R, represents a simple measure of the risk of policy failure. It is defined as the fraction of 1.5˚C runs that result in cumulative emissions closer to the 2.0˚C best guess prediction than the 1.5˚C best guess prediction.

Table 6.4 Indicators for overlap (O) and risk of policy failure (R)

R = Fraction of 1.5˚C runs that result in cumulative emissions closer to 2˚C than 1.5˚C. **O** = Fraction of runs in the area of the histogram covered by both the 1.5˚C simulations and the 2˚C simulations (i.e. the shaded turquoise area). R = 0.15 O = 0.61

The value of O of 0.61 confirms that the overlap between the 1.5C runs and the 2C runs is large: 61% of the runs for the two scenarios fall within this region. Of course, O, would be smaller if we compared impacts of policies designed to reach targets that were further apart. The reason why there is a much bigger overlap between emissions in the two policy scenarios than between emissions in the baseline and the policy scenarios is because the distance between the targets in the first case is much smaller. The spread in emissions in all three cases is actually very similar (as can be seen from the standard deviations in Table 6.3). At the same time, however, and as already noted, the difference between 1.5˚C and 2˚C is significant because of the large environmental impacts.

Despite the big overlap, the risk of policy failure as defined by R is relatively small (0.15). In fact, this points us to an interesting observation regarding the three distributions: they are all left skewed (i.e. the mean is higher than the median). The mean in all three cases also lie below the best guess predictions. Since input parameter distributions are completely symmetric, this indicates a non-linear response in FTT:Power to changes in core parameter values. If higher and lower deviations from the best guess parameter values are equally likely (as they are in the ±50% distributions used in this Chapter), this means that the best guess FTT:Power predictions provide conservative estimates of the emissions reductions potentials of policies.

In summary

1. Predictions of power sector emissions produced by FTT:Power are sensitive to core parameter values.
2. Policies, however, lead to significant emissions reductions compared to the baseline, even if we take into account variability of core parameters. In other words, low-carbon technology diffusion will not, on its own, be enough to reach the Paris target.
3. At the same time, the significant overlap in cumulative emissions in the 1.5C and 2C MCA runs indicate that the certainty of best guess FTT:Power predictions of the impact of policies on power sector emissions is limited. Recall that the analysis conducted her includes only a sub-set of FTT:Power parameters. Given the environmental significance of even small differences in emissions, such uncertainties should, according to the argument in Chapter 4, not be taken lightly.

So far, nothing has been said about which of the parameters have the largest impact on FTT:Power predictions. This is what the next section looks at.

6.5.3 Correlation analysis – most influential parameters

Pearson's correlation coefficient, which measures the strength and sign of the linear relationship between two variables, can be used as a simple measure to identify the strength of influence of parameters on emissions[81].

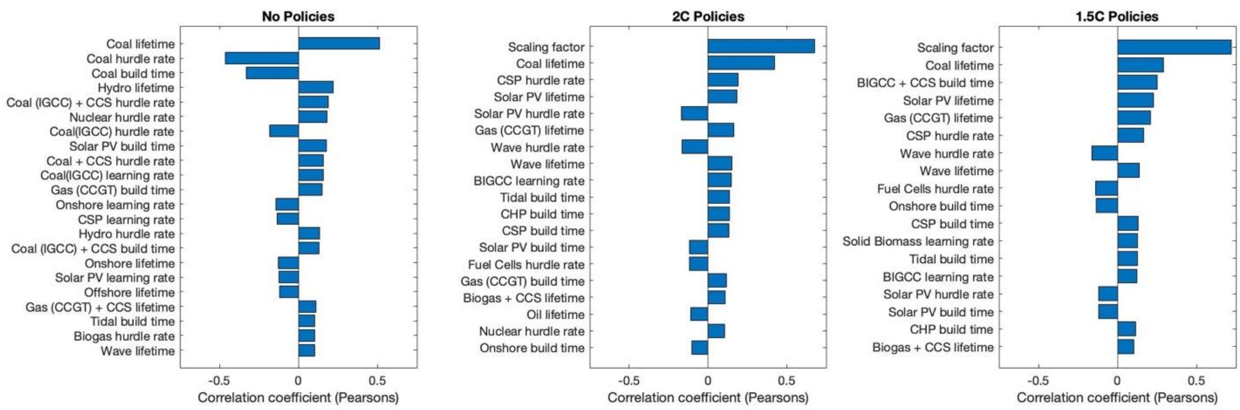


Figure 6.12 Parameters ranked according to their correlation coefficients (Pearson's) in each of the three Policy scenarios. Parameters with correlation coefficients less than 0.10 (and larger than -0.10) are not included.

[81] Note that this does not capture non-linear relationships

Figure 6.12 shows the parameters with the highest correlation with cumulative emissions in each of the three policy scenarios ranked according to the value of the coefficients. The figure shows all parameters with correlation coefficients larger than 0.1 and smaller than -0.1. The first thing to note is that only a handful of parameters in each policy scenario show a strong correlation. This is generally the case when performing sensitivity analyses, i.e. only a few inputs tend to substantially influence the value of a particular output (Loucks et al., 2005). Second, the parameters with the largest influence on cumulative emissions vary from one scenario to another. In the baseline, τ_{Coal} shows the strongest linear correlation, with a coefficient of 0.51. This is followed by r_{Coal} at -0.46, t_{Coal} at -0.33, and τ_{Hydro} at 0.22. For the 2C policy set, the scaling factor, K, shows the strongest correlation with emissions, with a coefficient of 0.68. This is followed by τ_{Coal} at 0.43. For the 1.5C policy set, the scaling factor is even more significant, with a correlation coefficient of 0.72. After this, four parameters have correlation coefficients in a similar region: τ_{Coal} at 0.29, $t_{BIGCC+CCS}$ at 0.25, $\tau_{Solar\ PV}$ at 0.23 and τ_{Gas} at 0.21.

Overall, parameters for coal (including the lifetime, build time, and investor discount rate) have a large influence on emissions in all scenarios. This is not surprising given that coal has the highest emissions coefficient of all the in FTT:Power. This means that the market share of coal, which again is particularly affected by coal-specific parameters, is a strong determinant of emissions. The parameter showing the strongest overall correlation, however, is the scaling factor. This is also not surprising given that this is a global parameter, meaning that it can have a large effect on the rate of diffusion overall. At the same time, it is worth noting that this parameter does not even show up among the most influential parameters in the baseline. This can be explained in the following manner: because the entire shares equation is multiplied by K, it means that K affects the rate of change of all technology shares simultaneously. Thus, when low-carbon technology diffusion is already set in motion by policies that render these technologies preferable to unabated fossil fuel technologies, an increase in the value of K will simply speed up this process. However, if technology diffusion is headed in a more neutral direction, making the power sector neither cleaner nor dirtier, the value of K, although it will also speed everything up in this case, will not have a significant impact on emissions. Thus, it seems, K only influences emissions when an increase in low-carbon deployment is already "set in motion" by policies. Clearly then, the value of K is key to FTT:Power predictions of the impacts of stringent climate policies. If the value of this parameter cannot be determined with a high degree of accuracy, predictions will be uncertain. In addition to this, renewable technology parameters also become more influential when policies are in place. For the 1.5C policy set, CCS technology parameters also start influencing emissions more. This makes sense given the importance of these technologies for reaching low levels of emissions.

The group of parameters with the least significant influence on emissions in FTT:Power are the learning rates. Given the well-known influence of learning on mitigation costs in IAMs, it is somewhat surprising to see that learning parameters have such a limited impact on emissions. This might be explained in part by the limited time period explored in FTT (2013-2050) compared to most IAMs (which typically run to 2100). Recall that the logistic equation, which gives rise to the S-shaped diffusion curves, implies that it takes time for new technologies, such as wind and solar PV, to grow. Because learning takes time, it might simply be that FTT:Power doesn't run long enough to capture the effects of learning on technology deployment. In comparison, changes to the other parameters have instantaneous effects on the competitiveness of technologies and the rates at which they can replace other technologies. This allows new technologies to reach larger market shares earlier on in the simulation period, which again has an impact on their ability to grow even more later on.

Lastly, it is worth noting that the MCA runs do not take into account the impacts of changing electricity prices (as a result of changing parameter values) on electricity demand, which again affects emissions. This means that MCA runs with parameter values that lead to a reduction in electricity prices will tend to underestimate the impacts on emissions and that MCA runs with parameter values that lead to an increase in electricity prices will overestimate emissions. Such price-demand feedbacks are included in the full E3ME-FTT model. As already noted, however, due to computational constraints, it is not feasible to perform the GSA that is conducted in this chapter for the full model[82].

Scatter plots for the parameters with the highest correlation coefficients are shown in figure 6.13. The scaling factor shows a clear linear relationship. In the baseline we see that the variance in cumulative emissions appears to be higher for low values of τ_{Coal} than for high values. This might be explained as follows: when the lifetime of coal is long, it is very difficult (or impossible) for other technologies to replace coal. This means that it is very difficult (or impossible) for cumulative emissions to reach low levels. The coal simply stays in the system for too long. For short coal lifetimes, we see instances of both high and low cumulative emissions. This implies that emissions in these cases depend more on other parameters. We see the opposite trends for t_{Coal} and r_{Coal}, as expected given they both have the opposite impact on the rate of change of the share of coal compared to τ_{Coal} (whereas a longer lifetime increases the growth rate of coal shares, a longer build time and a higher discount rate decrease the growth rate).

[82] While 200 runs can be done in just under two hours using FTT, it would take 100 hours using the full model.

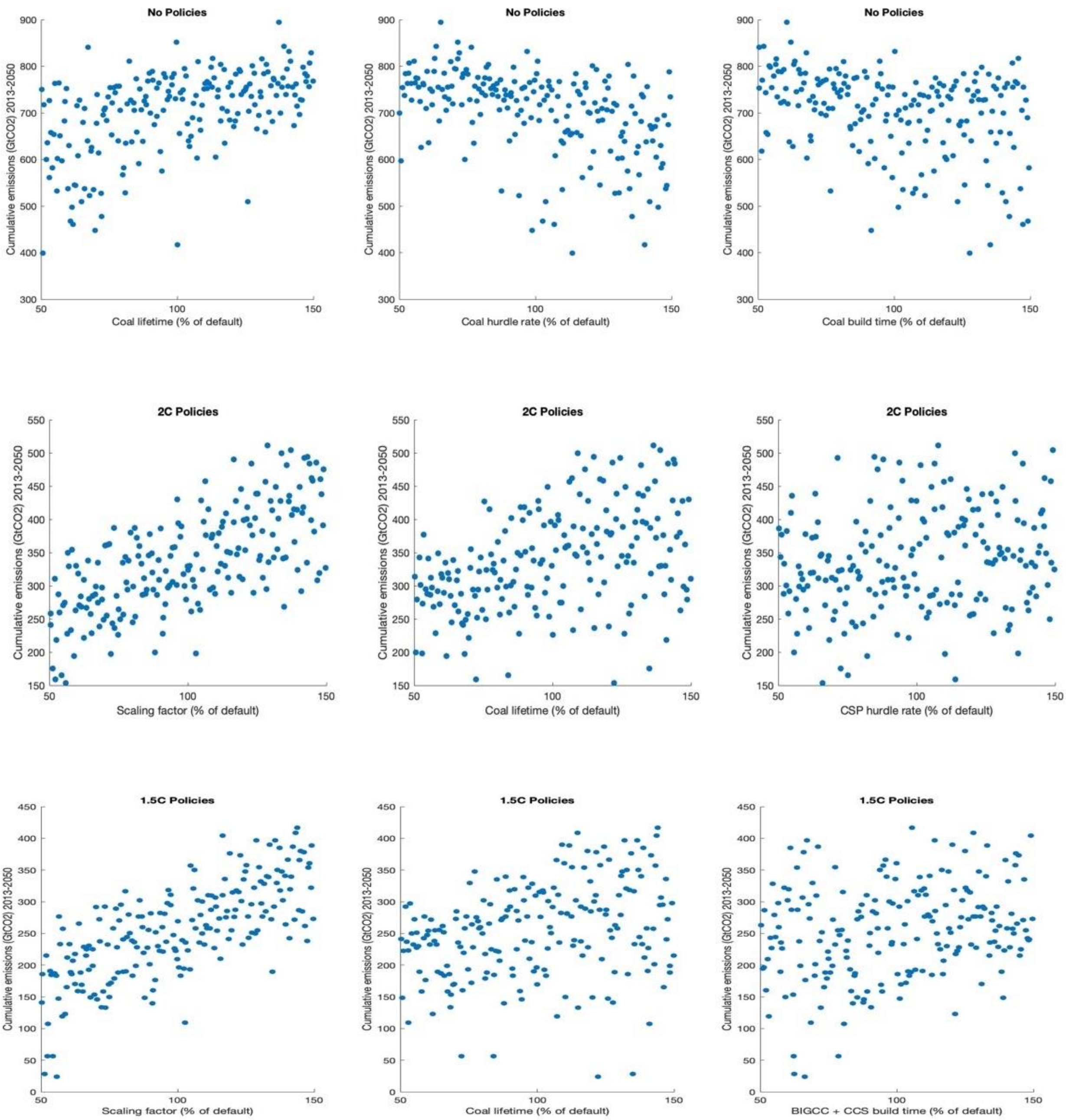


Figure 6.13 Scatter plots showing the relationships between cumulative emissions and the three parameters with the highest correlation coefficients for each of the scenarios.

In summary

1. The scaling factor is the single most influential parameter.
2. Of the technology-specific parameters, the parameters for coal have the largest influence.
3. The parameters with the least influence are the learning rates.

Given that FTT:Power includes 24 different technologies, we might not expect any single parameter value to have a large impact on emissions. Aside from the scaling factor, that is also what the correlation analysis shows. Although we have identified a handful of technology parameters (for coal in particular) that appear to have a significant impact on emissions, the runs in which emissions deviate significantly from the best guess predictions will be characterised by particular *combinations* of parameter values which on their own would not have very strong impacts on emissions. For instance, an MCA run that combines high build times for low-carbon technologies with low build times for fossil fuel technologies would lead to high emissions. A GSA is particularly useful in this regard because it automatically captures the likelihood of such events (given the specified input distributions). In other words, the distributions of outputs obtained from the MCA runs, which are shown in figures throughout this chapter, also take such possibilities into account.

Overall, the sensitivity analysis conducted in this chapter has shown that the uncertainty of FTT:Power predictions are large and that the scaling factor has the largest influence on results.

6.6 Levers of change

A sensitivity analysis can be used for more than just determining the robustness of predictions to the uncertainty of model parameters. It can also be used to identify additional levers of change. This section explains how this might be done.

Parameters differ along two important dimensions. First, in how sensitive they are. Secondly, in how much control we[83] have over them. When parameters are not sensitive, our degree of control matters little for outcomes. When parameters are sensitive, however, our degree of control becomes important; if we can influence parameter values, this can be used to push outcomes in a desired direction. Essentially, if we can influence the value of a parameter, that parameter starts to become a *control* or *decision variable*, not just a parameter to be estimated. A sensitivity analysis, at its core, shows how model outputs depend on model inputs. Whether the model inputs are considered *parameters* or *decision (*or *control) variables*, however, depends on the person who is using the model (Morgan & Henrion, 1990). The influence that a firm exerts is different from the influence that a national government exerts. Since E3ME-FTT has been used to inform policymaking at both European and national levels, it makes sense to consider policymakers at these two levels.

If policymakers can influence the value of a parameter, sensitivity to that parameter might be a good thing because it represents an additional lever of change. Lack of influence, on the other hand, when a parameter is sensitive, implies an increase in the uncertainty of predictions, which might imply

[83] Who “we” is referring to will depend on who is using the model predictions – see the discussion that follows.

increased risks[84]. Figure 6.14 places the five parameters analysed in this chapter according to the two dimensions, sensitivity and degree of control.

Sensitive	K			
		t_i	t_i	r_i
Not sensitive	LR_i			
	Low control			High control

Figure 6.14 Sensitivity and degree of control over parameters analysed in this chapter.

Policymakers have little control over learning rates (or at least limited knowledge of how to control them, which for all practical purposes is the same). Given their low sensitivity, however, this does not (according to FTT:Power) have a big impact on emissions reductions in the power sector[85]. Build times have some impact on emissions. There are elements of the build times that are outside policymakers' control, including supply chains, construction times, and availability of labour. If we include planning and permissions into build times, however, as the acceptance of new energy technologies improve and institutions gain experience (a process that might be labelled 'institutional learning') the build times might decrease for new technologies. Related, it is also conceivable that governments could decide to prioritise low-carbon technology roll-out and thus speed up build times. In this sense, policymakers might exert some influence over build times as represented in FTT:Power.

Lifetimes are similar to build times, but perhaps more straight-forward in terms of control. Just as lifetimes for power plants are often extended beyond what was initially intended, lifetimes can also be shortened by early retirement. In this sense, policymakers might also exert some influence over lifetimes as represented in FTT:Power.

Investor discount rates are determined by the cost of capital and the risks associated with an investment as well as the extent to which investors are far- or short-sighted (see e.g. Peters et al. (2011)). This again is influenced by financing structures and institutions (Egli et al., 2018), which again are influenced by policies. For instance, different institutional structures appear to make German companies more far-sighted than English companies. There is currently a suggestion to include workers on boards in the UK

[84] Assuming that risk is defined as the likelihood times the consequences of an event, an increase in uncertainty will generally imply an increase in risks.

[85] As already noted, however, learning rates might still be important in the longer run. FTT:Power is limited in that it only provides predictions up to 2050. Additionally, learning rates will have an impact on costs, and therefore be important for climate policy in general.

(Pratley, 2018) with the goal being that this would make decision making more far-sighted (following the German model). In addition to this, governments can also influence the risks of different investments directly by issuing insurances and guarantees. Thus, policymakers have an influence over investor discount rates.

In short, policymakers may thus exert some influence over the build times, lifetimes, and investor discount rates that are represented in FTT:Power as input parameters. FTT:Power tells us that the relative values of these parameters for different technologies impact the diffusion of low-carbon technologies. The point here is that this can be considered an insight (if one believes FTT:Power to provide a good description of power sector dynamics[86]) for policymaking: in addition to the design of carbon pricing, subsidies, feed-in tariffs, and kick-start policies – which are all considered policy variables in FTT:Power – policymakers might also want to think about the other ways in which policies might have an impact on low-carbon technology diffusion. This is one of the ways in which FTT:Power might be used to provide insights – not just numbers[87] – for policymaking.

6.7 Conclusion

In this chapter, a global sensitivity analysis (GSA) has been conducted in order to identify the parameters in the FTT shares equation that have the largest impact on FTT:Power predictions and to provide a first conservative estimate of the uncertainty of FTT:Power predictions. This was done by using a Monte Carlo analysis with LHS and assuming that parameter values are uniformly and independently distributed around ±50% of their default values. A GSA was chosen in order to provide a more accurate picture of how the uncertainty of key parameters in the FTT model affect the value of emissions and technology deployment in FTT:Power, which is a complex non-linear model, than what a one-factor-at-a-time approach would have done.

The results of the analysis show three things. First, if we assume that FTT:Power provides a sufficiently accurate description of power system dynamics and the ways in which these dynamics are impacted by policies, the results show that the diffusion of low-carbon technologies on its own, even when we allow for optimistic assumptions regarding the values of technology parameters, is unlikely to lead to transformation pathways compatible with the Paris target in the power sector. In other words, policies are necessary to reduce emissions.

---

[86] It should be noted that the fact that build times, lifetimes, and investor discount rates affect technology diffusion is really an assumption in FTT that stems from diffusion theory (see e.g. Mercure (2015)).

[87] Harking back to the often repeated statement by Peace and Weyant (Peace & Weyant, 2008).

Second, the sensitivity analysis shows that the scaling factor – a constant that represents the time it takes to achieve a full turnover of technologies – has a large influence on FTT:Power predictions of the impacts of policies on technology deployment and emissions in the power sector. The fact that the correct value of this parameter is deeply uncertain poses a serious issue for the accuracy of FTT:Power predictions.

Third, the analysis suggests that policymakers might be able to exert some influence over some of the parameters that, according to FTT:Power, determine the rate of technology deployment. In other words, some of the parameters investigated in this chapter might be viewed not simply as parameters whose values need to be determined, but as decision (or control) variables. Based on an assessment not only of the sensitivity of different parameters, but also of policymakers degree of control over them, it appears policymakers could speed up the diffusion of low-carbon technologies by reducing the risks associated with low-carbon investments, and by prioritising the roll-out of low-carbon projects over fossil fuel projects. Given the influence of coal parameters on results, not letting coal power plants run beyond their lifetimes (or retiring coal power plants early), would also help by leaving space for other, less emissions intensive electricity generating technologies to diffuse. Investments in dirty technologies can also be made less desirable by influencing the risks associated with different technologies (several insurance companies are, for instance, starting to decline to insure coal investments).

A number of caveats are important to note when interpreting the results of the analysis conducted in this chapter. First, the GSA does not include all FTT:Power parameters. Although all the parameters that define the FTT core equation, the shares equation, were included, the fact that only a sub-set of all the parameters were included means that the GSA does not necessarily identify the parameter that influence FTT:Power predictions the most. Future analyses should therefore expand the number of parameters to include those that are left out here (a list of the remaining parameters is provided in Appendix X). Due to the way that the model is implemented in MATLAB (with multiple functions that are called in different parts of different programs), this would require some work. Second, parameter values are treated as independent of each other in the current analysis. This means that some combinations of input parameters, all of which are treated equally in the MCA in this chapter, are more likely than others. This again will have an impact on the distribution of outputs. Third, the results say nothing about costs. FTT computes the impacts of policies on technology deployment and emissions. It does not tell us anything about the costs of different scenarios. While the sensitivity analysis might give the impression that the lower emissions are, the better, it is possible that for some of the low emissions runs, costs are also higher. This is clearly important to policymakers when deciding what policies and additional measures to implement. To assess macroeconomic costs, E3ME would have to be included in the analysis. This, however, would not be feasible given the large number of runs required for the MCA. One alternative could therefore be to develop an FTT:Power version that uses price elasticities to approximate macroeconomic impacts. Another alternative could be to use estimates of the costs of generating electricity in different regions in FTT:Power to develop an aggregate proxy for the cost of generating electricity. All of these possibilities represent non-trivial options, which may be pursued in future uncertainty analyses of FTT:Power. The goals of the analysis conducted in this chapter were to provide a first conservative estimate of the uncertainty of FTT:Power predictions and to identify the parameters in the FTT core equation that have the biggest influence on FTT:Power predictions.

If one wanted to claim that the distributions of outputs from the MCA runs reflect the "true" uncertainty of FTT:Power predictions – not just their sensitivity to the parameters that are included in the analysis or a conservative first estimate – one would have to not only include all (uncertain) model parameters, but also make sure that the parameter distributions reflect the "true" uncertainty of the parameter values. This, by all accounts, is no small endeavour. In fact, the GSA conducted in this chapter already shows that the parameter that influences FTT:Power predictions the most is characterised by deep uncertainty; It is not clear how one would determine the likelihood with which this parameter would take on different values. Nonetheless, future uncertainty analyses of FTT:Power may benefit from basing parameter distributions on more comprehensive literature reviews or expert elicitations (such as is done by e.g. Chan and Anadon (2016) and Bosetti et al. (2015)). Given that any attempt to determine the "true" uncertainty of parameter values in FTT:Power will require time and effort, the GSA conducted in this chapter provides valuable information regarding what parameters to focus on: the overall scaling factor and coal parameters (the lifetime, build time, and investor discount rate). At the same time, unless one also assesses the structural uncertainty of FTT:Power, one should still be careful to interpret the results of a parametric uncertainty analysis of FTT:Power as reflective of the "true" uncertainty of the impacts of policies on technology deployment and emissions in the power sector.

Despite the above limitations, the GSA conducted in this chapter shows that the impacts of policies on technology deployment and emissions predicted by FTT:Power are sensitive to uncertain parameter values. In particular the sensitivity of FTT:Power predictions to the overall scaling factor, whose "true" value is deeply uncertain suggests that FTT:Power is not able to predict the impacts of policies on technology deployment and therefore emissions in the power sector to a high degree of accuracy. This is true even in the near-term, where we might expect predictions to fare better. Although the distributions of outputs computed in this chapter should not be interpreted as reflecting the "true" uncertainty of the respective outputs, there are reasons to believe that the distributions represent conservative estimates of the uncertainty of FTT:Power outputs. First, the comparison of parameter distributions used in this chapter with the ranges found in the literature (shown in Appendix D) suggests that the ±50% ranges are within the bounds of what is thought to be possible. Second, even if some of the ranges are slightly wider than what is found in the (brief) review shown in Appendix D, the analysis only includes a sub-set of all FTT:Power parameters (the full list of FTT:Power parameters is provided in Appendix X). If more parameters were included in the GSA, the distribution of output values could only increase. The analysis conducted in this chapter shows that the range of outputs that results from varying even just a sub-set of FTT:Power parameters within what is seen to reflect a plausible range is large. Third, the analysis conducted in this chapter completely ignores structural uncertainty. According to Morgan and Henrion (1990), experienced analysts tend to argue that structural uncertainty is more likely to have a significant impact on results than parametric uncertainty. Future work is needed to investigate, among other things, the extent to which the species competition analogy used to derive the core equation in FTT:Power provides a good analogy for the dynamics of technological change in the power sector. In part because FTT is a new and different model, relatively little attention has so far been paid to verifying (to the extent this is possible) the structural assumptions. The assumed decision making of investors based on LCOEs could also be better backed up. Based on all of the above, it is reasonable to argue that the output distributions presented in this chapter represent conservative estimates of the uncertainty of FTT:Power predictions.

The goal of FTT is to predict the impacts of energy and climate policies on technology deployment and emissions. The fundamental uncertainty of the overall scaling factor in FTT:Power and its influence on results, on its own, challenges the reliability of FTT:Power best guess predictions. As argued in Chapter 4, the importance of the uncertainty of the impacts of policies depend on the consequences of being wrong. In climate policy, the environmental consequences of acting too late or too little are generally irreversible, and potentially dire[88]. If parameters that have a large influence on results are characterised by fundamental uncertainty, providing policymakers with best guess predictions in this domain is not defensible. When the uncertainty is important, it is more appropriate to provide policymakers with ranges of results contingent on uncertain parameter values. This is done, for example, by the Interagency Working Group in the US Government when reporting the social cost of carbon, which is largely dependent on the discount rate (Interagency Working Group on Social Cost of Carbon, 2010)[89].

Chapter 5 showed how FTT is seen by Mercure et al. to offer a "a more realistic modelling approach" (2016, p. 102), which in turn is seen to lead to better predictions of the impacts of policies on technology deployment and emissions compared to what you get with ESOMs. While ESOMs are criticised for their reliance on neoclassical assumptions (see Chapter 5), FTT is seen to offer an improved description

[88] Noting that the consequences of being wrong with respect to FTT predictions is different from the consequences of being wrong with regards to cost estimates reported in IPCC reports discussed in Chapter 4. Environmental consequences of climate change, however, remain significant.

[89] Although the interagency working group uses only three discrete values of the discount rate, the point is the same: when key assumptions are uncertain, it is more appropriate to present a range of values than single best guesses.

of technological change based on the theory of technology diffusion. Additionally, Chapter 5 showed that the rate of technology deployment in ESOMs is determined to a large extent by exogenous constraints. Does FTT fare any better? The analysis in this chapter shows that the use of diffusion theory to endogenously derive deployment rates in FTT:Power does not avoid the dependency of results on exogenous assumptions. Nor does the value of the scaling factor appear to be any less uncertain or any easier to verify than the value of the maximum technology deployment rates that are used in ESOMs. In fact, both the scaling factor and the maximum deployment constraints express our beliefs about the rate of technological change in the future. Thus, it appears, predictions generated by both ESOMs and FTT:Power depend on similar unknowns.

Overall, the analysis conducted in this and the previous chapter thus appears to confirm the widely held view that technological change is inherently uncertain and difficult to predict. In many ways, it should not be surprising that future technology deployment is mired in uncertainty. The question is what we do with this uncertainty. While the uncertainty does not negate the insight that policies are necessary to reduce emissions, it does raise questions regarding the use of deterministic best guess predictions of policy impacts to inform decision making. If the impacts of policies on technology deployment in FTT:Power are uncertain, they will also be uncertain in the full E3ME-FTT model. The latter model is frequently used by the European Commission to inform energy and climate policy as well as long-term strategies for emissions reductions (e.g. European Commission (2018)). If E3ME-FTT predictions of policy impacts are inaccurate, resulting policies might be poorly informed. Given the limited time available to reduce emissions sufficiently to avoid "dangerous" climate change, the consequences of implementing policies that turn out to be insufficient might be significant environmental harm.

In summary, although transformation pathways generated by ESOMs are determined to some extent by exogenous assumptions regarding maximum deployment rates, transformation pathways generated by FTT:Power are to some extent determined by the scaling factor. In order to claim that FTT:Power provides predictions that are better than those provided by ESOMs, the value of the overall scaling parameter in FTT:Power needs to be better grounded in theory and evidence than what it currently is. In addition to identifying the scaling factor as a key source of uncertainty – which should be addressed in future modelling work – the analysis conducted in this chapter provides the first step towards using energy system models in a way that acknowledges and communicates to policymakers the important uncertainties that are present in this domain of research.

7 Conclusion

This thesis has identified and taken first steps towards assessing three challenges associated with the insights that can be drawn from IAMs: i) the importance of model independence for the robustness of insights that can be drawn from IAM ensembles and the lack thereof between many IAMs in AR5 (chapters 2-3), ii) the importance of the uncertainty of the cost of mitigation and the failure of the AR5 IAM ensemble to capture the uncertainties associated with this measure in the literature (Chapter 4), and iii) the dependence of FTT predictions on a deeply uncertain scaling parameter and the resulting uncertainty of best guess predictions (chapters 5-6).

Sections 7.1 and 7.2 provide detailed summaries of each chapter, list the contributions to the literature, and suggestions for future research. Section 7.3 summarises the main findings and offers concluding remarks on the role of diversity and its implications for IAM research.

7.1 Robustness and uncertainty in IAM ensembles

Chapter 2 argued that IAM independence is an important but neglected topic in AR5. IAM independence, i.e. independence of model choices and assumptions, is important because it is a prerequisite for drawing robust insights from IAM ensembles. If the IAMs in an IAM ensemble are not independent, we cannot know whether agreement in outputs is a sign of robustness or a consequence of shared model choices and assumptions. To assess IAM dependencies, Chapter 2 developed a method for constructing a model family tree based on model links found in model documentation. This method was used to identify likely model dependencies among IAMs in AR5. The analysis showed that the 14 most influential IAMs in AR5, which together are responsible for 90% of the scenarios in the AR5 scenario ensemble, form three branches, the largest of which is the MESSAGE/MERGE branch (consisting of MERGE, MESSAGE-MACRO, MERGE-ETL, REMIND, WITCH, and BET). The IAMs in this branch are responsible for about half of the scenarios in the AR5 scenario ensemble. The analysis of the model documentation furthermore indicated that the evolution of IAMs has been driven by a growing set of policy questions that has incentivised a continuous increase in the level of detail and scope of IAMs. By considering not only what was stated, but also what was not stated in the model documentation, Chapter 2 found that new IAMs rarely challenge the model choices and assumptions that are made in existing IAMs. All of this suggests an expansion of existing IAM approaches rather than an increasing diversity of approaches.

The findings in Chapter 2 indicate that the lack of independence between IAMs in AR5 might weaken the robustness of AR5 IAM results. While several authors have criticized IAMs for a lack of transparency regarding input assumptions and their impacts on results (e.g. Schneider (1997), Rosen

(2015), “IAM helpful or not?” [editorial] (2015)), the analysis conducted in Chapter 2 highlights the need also for transparency around shared model choices and assumptions and their impacts on ensemble results. Although more research is needed to determine the impacts of the lack of independence between IAMs in AR5 on AR5 ensemble results, Chapter 2 argued, model dependencies should be acknowledged and communicated in the outer layers of IPCC reports (i.e. in SPMs and the synthesis report).

Chapter 3 showed that there is a large overlap between the “key differences in model structure” discussed in AR5, the main model frameworks that underpin AR5 IAMs (optimal growth theory, CGE modelling, and ESOMs), and the branches in the model family tree constructed in Chapter 2. The overlap between the model family tree and the key structural differences indicates that most of the model links in Chapter 2 either directly capture or are closely related to similarities in model structure. This is not surprising given how the model links in Chapter 2 reflect model combinations, model versions, structural similarities, and shared model components, all of which are likely to either require or give rise to structural similarities. Given their status as *key* differences, it is reassuring to see that the method developed in Chapter 2 capture many of these.

At the same time, Chapter 3 also showed, the method developed in Chapter 2 captures both model dependencies that go across key structural differences (e.g. between GCAM and Phoenix) and model independencies that are not visible based on key structural differences (e.g. between IMACLIM and Phoenix). Thus, the information contained in AR5 (and in similar IAM classifications) is not sufficient to capture all IAM dependencies. This is not surprising given the many sources of model dependencies (some of which were discussed in Chapter 2 for climate models), and it demonstrates the value of the method developed in Chapter 2, which is able to capture some of the social scientific network links (e.g. those related to institutions) that are also at play.

Chapter 4 started from the observation that all the estimates of the cost of mitigation generated by the IAMs in AR5 are net positive. According to the literature, however, the cost of mitigation could be both net positive and net negative. Experts disagree. This implies that the scenarios in the AR5 ensemble do not reflect the full range of uncertainties regarding the cost of mitigation. Based on a debate on values in science in philosophy, Chapter 4 developed an argument for why this uncertainty is *important*. In short, due to the global scale and irreversible nature of the climate change issue and the fact that net negative cost results could have led to earlier and stronger action on climate change, being wrong about the sign of the cost of mitigation (i.e. failing to capture the possibility of net negative costs when the cost in reality is net negative) could have large negative consequences. This means that the uncertainty regarding the cost of mitigation is *important*. The failure of the AR5 IAM ensemble to reflect the possibility of net negative costs is therefore problematic.

When it comes to the reasons why the AR5 IAM ensemble contains only net positive cost results, Chapter 4 found, based on a review of the AR5 scenario publications, that, although general equilibrium IAMs (which are responsible for all the cost estimates in AR5) *can* be modified to reflect the possibility of net negative costs, only two IAMs in AR5 (IMACLIM and, to a lesser extent, WITCH) incorporate mechanism that, according to the literature reviewed in Chapter 4, typically contribute to such results. Additionally, Chapter 4 showed, the model intercomparison studies that are responsible for the majority (around 95%) of the scenarios in the AR5 ensemble focused on aspects that can only increase the cost of mitigation. Based on this, Chapter 4 concluded, there is reason to believe that the AR5 IAM ensemble might be biased towards net positive mitigation costs.

As with any analyses, there is a trade-off between breadth and depth. One of the key challenges associated with studying IAM ensembles *as ensembles* is that the number and complexity of IAMs in such ensembles preclude detailed comparisons of individual IAMs. This means that the analysis of IAM ensembles can appear shallow. Generalisations are susceptible to counterexamples, and results tend to be indicative rather than conclusive. This is true to some degree also for the results of chapters 2-4 in this thesis. The difficulty of studying IAM ensembles in fact highlights one of the very issues associated with drawing insights using IAM ensembles: the difficulty with which results can be evaluated. Several authors have long pointed out that results generated by individual IAMs are difficult to interpret because they are products of "black boxes" (Funtowicz & Ravetz, 1990; Keepin & Wynne, 1984; Stanton et al., 2008). This is even more true for results generated by IAM ensembles.

At the same time, by not studying IAM ensembles as ensembles, we might fail to discover some of the key challenges and to understand the conditions that are necessary for obtaining robust and reliable insights. Given the central role of IAMs as tools for assessing how to reach the Paris climate target, the ability to perform independent reviews is important. It would therefore be advantageous if these could be performed, at least in part, without the involvement of modellers. This will be easier if IAM researchers clearly communicated the model choices and assumptions that determine results and their level of confidence and reasons for making said model choices and assumptions.

The arguments presented and the analysis conducted in chapters 2-4, which focused on the AR5 IAM ensemble, are relevant both for the upcoming IPCC AR6 and for the wider IAM community, which is increasingly turning towards model intercomparison projects (e.g. NAVIGATE[90], ENGAGE[91], COMMIT[92], and PARIS REINFORCE[93]) that generate large ensembles of IAM results.

The findings also speak to current debates about whether (Anderson & Jewell, 2019) and how (Grant et al., 2020; Hausfather & Peters, 2020; Mccollum et al., 2020) IAMs should be used. A diversity of model choices and assumptions, this thesis concludes, is key to ensure robust insights (Chapter 2) and to make sure important uncertainties are captured (chapters 4 and 6). This call for diversity is closely aligned with the call from McCollum et al. (2020) for a more systematic exploration of extremes. While McCollum et al. focus on the exploration of extremes by varying structural and parametric assumptions in individual models (which is more in line with what is done in chapters 5 and 6), this thesis argues that we also need a diversity of modelling approaches to properly explore and understand the uncertainties associated with IAM ensemble results.

In addition to this, Chapter 4 also speaks to the debate on "neoclassical" (or "mainstream") versus "heterodox" (or "non-mainstream") approaches to economics that much of the literature on IAMs is situated within. For instance, Scrieciu (2007) warns against the 'inherent dangers' of using CGE models as a single integrated framework for sustainability impact assessment, Barker et al. (2012) argue for a Post Keynesian 'new economics' approach to climate policy, and Farmer et al. (2015) argue that a 'new wave of models' need to be developed to tackle current inadequacies in climate economics. While much of this literature argues against "mainstream" approaches (and advocates for various "non-mainstream" approaches), this thesis argues that all reasonable model choices and assumptions should be included in IPCC reports, which are meant to assess the strength of and uncertainties in scientific understanding related to climate change impacts, mitigation, and adaptation. Because the tails of the IAM output distributions can be important, Chapter 4 argued, non-conventional IAMs, as well as "optimistic" and "pessimistic" assumptions, serve an important purpose and should not be excluded from IPCC reports. An explicit goal of maximizing the diversity of IAMs included in IPCC reports could increase the robustness of findings or, alternatively, help illuminate important uncertainties.

Because little research has so far been done on IAM ensembles *as ensembles*, much of this thesis charts new territory and some questions raised are left for future research. In particular, understanding the processes by which scenarios are included or excluded in IPCC reports represents an important area of

[90] https://navigate-h2020.eu/

[91] http://www.engage-climate.org/

[92] https://themasites.pbl.nl/commit/

[93] https://www.paris-reinforce.eu/

future research. Further research on IAM ensembles could take at least two different directions. One possibility is to move into the sociology of science to obtain a better understanding of why and how assumptions and model choices are made and how IAMs end up being included in IPCC reports. What might lead to the exclusion of certain approaches or assumptions, and, as a result, a lack of diversity in the IAMs that dominate not only IPCC reports but the IAM literature in general? The analysis conducted in Chapter 4 revealed that model intercomparison studies, which are responsible for about 95% of the scenarios in the AR5 scenario ensemble, exert a large influence on AR5 results. What might limit the inclusion of diverse perspectives in such studies and in the IPCC? What happened to IAMs that were previously included in IPCC reports (such as E3MG) that no longer are? Related, should the IPCC continue to rely purely on scenarios published in the peer reviewed literature? This requirement imposes limitations on the scenarios that can be included and introduces time delays (between research and publication), academic incentives, and hierarchies (Wright, 2018). There is a large and growing literature on perverse academic incentives (e.g. Edwards and Roy (2017)). But what would the alternative to peer-review be? Research in the direction of sociology of science could also examine the influence of IAM results on climate policy and action. To what extent do IAM results influence policymakers, private actors, and the public debate? How are IAM results interpreted, and are they trusted? Does it matter if IAM results are inaccurate or misleading? What IAM results are most important and why?

Another possibility is to go down a quantitative route focused on analysing existing IAM results. Can we infer anything about the diversity (or lack thereof) of model choices and assumptions from the spread or clustering of results in existing scenario databases? How have IPCC ensemble results changed over time? How have individual IAM results, such as estimates of the cost of mitigation, changed over time and how do they vary between model intercomparison studies and individual scenario publications? How would a different weighing of scenarios in ensembles affect reported averages and median values?

7.2 Reliability of IAM results

Chapter 5 presented several claims put forth by FTT modelers regarding the superiority of FTT relative to ESOMs. The chapter focused in particular on the claim by Mercure et al. (2014; 2016) that the endogenous derivation of technology deployment rates based on the theory of technology diffusion enables a more realistic depiction of the impacts of policies on future technology deployment. While the chapter agreed that ESOMs are unlikely to provide good predictions if the assumptions of perfect markets, rationality, and maximum technology deployment fail, it argues that we cannot claim that FTT predictions are any better without (at least) assessing the sensitivity of these predictions to key uncertain assumptions.

In order to do so, Chapter 6 conducted a global sensitivity analysis of the power sector sub-model of FTT (FTT:Power) based on Monte Carlo analysis and Latin Hypercube sampling with the goal of providing a first conservative estimate of the uncertainty of FTT:Power predictions and identifying the parameters in the FTT core equation (the shared equation) with the largest influence on results. This is the first time such an analysis has been conducted for any of the FTT models. Using uniform and independent distributions that span ±50% of default parameter values, the sensitivity of technology deployment and emissions in FTT:Power to investor discount rates, technology build times, technology lifetimes, learning rates, and the overall scaling factor – a parameter representing the time it takes to achieve a full turnover of technologies – was computed. Given that the analysis included only a sub-set of FTT:Power parameters and ignored structural uncertainty (and given that the ±50% ranges were shown to be close to ranges found in the literature) the results of the analysis conducted in Chapter 6 can be interpreted as a conservative estimate of the uncertainty of FTT:Power predictions.

According to the results, the impacts of policies predicted by FTT:Power are highly sensitive to the scaling factor, whose value is deeply uncertain. While this does not negate the result (if we accept the structural assumptions made in FTT) that policies are likely to be necessary to reduce emissions (a results that holds across a wide variety of parameter values), it does raise questions regarding the use of best guess FTT:Power predictions to inform policymakers about the impacts of policies. Given the importance of the uncertainty (as argued in Chapter 4), Chapter 6 argued, it is more appropriate to provide policymakers with ranges of results contingent on key parameter values than it is to provide them with best guess predictions when the latter are highly uncertain.

Overall, chapters 5 and 6 showed that, while the rates of technology deployment in ESOMs are determined partly by exogenous constraints, the rates of technology deployment in FTT:Power are determined to a large extent by the scaling factor. Thus, the use of diffusion theory to derive deployment rates in FTT:Power does not in itself circumvent the dependency of results on uncertain and debatable assumptions. The value of the scaling factor appears to be no more certain or any easier to verify than the values for the maximum technology deployment rates that are assumed in ESOMs. Moreover, the scaling parameter is not discussed (or mentioned) in FTT publications. The results of the GSA conducted in Chapter 6 thus contributes to a long list of studies going back to at least the 1980s (e.g. Keepin and Wynne (1984)) that have shown that energy system model outputs can be highly sensitive to uncertain assumptions. The results also lend support to the claims that large models often generate spurious detail (Funtowicz & Ravetz, 1990; Morgenstern, 1963).

Lastly, it is worth noting that multiple authors, in addition to Mercure et al. (2014; 2016), have criticized IAMs for an inadequate representation of policies and real-world processes, including innovation and diffusion (e.g. Farmer et al. (2015), Stern (2016), and Rosen and Guenther (2015)). Chapter 5 and 6 contribute to this literature by discussing the use of ESOMs as predictive tools and highlighting the dual role that ESOMs have played as tools for policy optimization (in “optimisation mode”) and policy evaluation (in “simulation mode”). This discussion is useful because the goal of the modelling matters for the importance of the noted inadequacies. If the goal is to identify optimal solutions (and those who use the models and their results are aware of this), the fact that models don’t capture real-world policies is less of a problem; knowing the least-cost solution can be useful for policymakers even if the models don’t say anything about how to reach that solution. If the goal, on the other hand, is to predict the impacts of policies, the inability of models to capture real-world policies can seriously diminish the value of the resulting predictions.

7.3 Diversity matters

Diversity in modelling approaches, or more specifically in model choices and assumptions, this thesis has argued, is important. There are two main reasons for this. First, diversity is crucial for our ability to obtain robust insights, and second, diversity is crucial for reflecting important uncertainties associated with IAM research.

When IAMs are not independent, diversity is limited. When this is the case, we cannot *know* whether agreement in results is a consequence of shared assumptions and model choices or a sign of robustness. In an ideal situation, the all plausible model choices and assumptions are captured by the IAMs in an ensemble. When IAMs agree, in this case, we would *know* that this agreement indicates a robust insight. When IAMs disagree, in this case, we would *know* that the result depends on model choices and assumptions that are either uncertain or open to disagreement. Both situations provide valuable insights: they tell us what we know and what we don’t know, what findings are certain and what findings are uncertain. Recall that “[a]n integral feature of IPCC reports is the communication of the strength of and uncertainties in scientific understanding underlying assessment findings” (IPCC, 2014b, p. 37). If IAM ensembles are based on only a narrow range of plausible model choices and assumptions (which is more likely to be the case when IAMs are not independent) we cannot conclude that agreement in results represent robust insights. On the contrary, if IAM ensembles fail to incorporate a diversity of approaches, results that in reality are uncertain (or open to disagreement among experts) might be *mistaken for* robust insights. The result that the cost of mitigation must be net positive might, according to the findings in this thesis, represents one such example.

When IAM ensembles are based on a more diverse set of approaches and assumptions, the results begin to approximate the “true” uncertainty of IAM research. This uncertainty can, as Chapter 4 argued, be important, especially for key results that matter to policymakers and negotiators such as the cost of mitigation. Chapter 6 also showed how single model predictions can become highly uncertain when ranges of plausible parameter values are considered. Uncertainty is more important when the consequences of being wrong are large. When it comes to climate policy, however, both economic and environmental consequences tend to be large simply due to the nature and the scale of the problem. The stakes are high.

While it is relatively easy to capture the uncertainty of parameter values in IAMs, IAM ensembles are unlikely to ever capture all plausible assumptions and model choices. It is not even clear what "all plausible assumptions and model choices" means or who would decide when they have been captured. There will always be debate and disagreement within the IAM community regarding what assumptions are plausible and what assumptions are justified (see e.g. the debate about the appropriate discount rate in CBA IAMs (Dietz et al., 2007; Nordhaus, 2007)). In addition to this, there are practical and computational limitations regarding what model choices and assumptions *can* be incorporated in IAMs. Nonetheless, diversity in IAM ensembles can still serve as a useful goal. As indicated in Chapter 2, there are no obvious signs that the drivers behind IAM development over time has incentivized a diversity of approaches.

Although using a diversity of approaches would better reflect uncertainties associated with IAM research, there are also downsides to an increased focus on uncertainty and the multitude of assumptions and model choices that, in theory, may be deemed plausible. Results might be more difficult to interpret and communicate. The uncertainty might, in some cases, overwhelm users of IAM research and leave them confused as to what to take from it. Uncertainty might also deter action, especially when there are large costs involved.

Ignoring uncertainties, however, also poses serious issues. The course of action to limit global warming is a political question. When outcomes are uncertain, and stakes are high, choices become even more value laden (Funtowicz & Ravetz, 1993). Do we go all in to avoid worst case outcomes, or do we try to minimise the cost of action? What is worse? Failing to reduce emissions sufficiently to limit global warming to "well below 2°C" or implementing policies that turn out to have large costs that are perhaps unnecessary for avoiding so-called "dangerous" (Schneider, 2001) climate change? While some might argue we should base decisions on expected utility, others might argue for strategies that primarily seek to avoid worst thinkable outcomes. These kinds of questions have become even more relevant with the debates around COVID-19 scenarios and strategies. What we do when faced with a potentially great threat is, in most people's opinion, not a question that can be answered by science alone (see e.g. Weinberg (1972) for an early account of this debate). While IAM researchers might agree with this, the general practice of leaving out uncertainties associated with IAM results might *unintentionally* lead to a premature narrowing of political deliberation (Beck & Krueger, 2016; Stirling, 2010). Some of the uncertainties, for instance with respect to the cost of mitigation and the effects of policies on technology deployment and emissions, are themselves relevant and important for decision makers. Not communicating uncertainties takes away the opportunity for decision makers and stakeholders to discuss the options they deem preferable *given* uncertainties, including hedging or precautionary strategies. An increased emphasis on uncertainty in IAM research can thus provide valuable inputs to the democratic debate regarding what to do about climate change and how to reach the Paris target (Stirling, 2010).

Does this mean that we should include all plausible options? Yes, this thesis argues, as long as these options are plausible and can be justified. In addition to the reasons for taking uncertainties seriously outlined above, three additional considerations are worth noting.

First, most IAM researchers are unlikely to have the area-specific knowledge required for making judgments about the uncertainty in every domain that is relevant to IAM research. While IAMs are good tools for assessing the consequences of making certain assumptions in a consistent manner, considering an enormous amount of data and interactions at a global scale, IAMs are not good tools for assessing the uncertainty of input assumptions such as learning rates for different technologies, the extent to which market failures beyond the climate externality are present, or the availability of CCS in the future. It would therefore be more transparent if IAMs were used to show what assumptions matter, while leaving for others the assessments of whether those assumptions are likely to hold. In short, IAMs can be very useful for identifying key assumptions. They are less useful for narrowing down the uncertainty of said assumptions.

Second, policymakers and the public can only learn to understand and cope with uncertain findings if researchers make an effort to communicate and explain the uncertainty. Not including the uncertainty of IAM results might also reduce the credibility of IAM research and diminish their value as inputs to climate policymaking in the long run. IAM researchers should therefore be more transparent when it comes to the uncertainties associated with their results. The role of research is not just narrowing down uncertainty, but to determine when uncertainty is high and when it is low.

Third, the *importance of the uncertainty* varies considerably. This thesis has argued that it depends on the risk of being wrong. When the risk is high, we should pay more attention to the uncertainty, hence it should be more carefully communicated. When the risk is low, we need not worry about uncertainty as much. Thus, IAM researchers do not necessarily have to worry about uncertainty all the time. Furthermore, far from all assumptions have a significant impact on IAM results (those that do can be determined via a sensitivity analysis). Thus, it will generally not be necessary to examine the uncertainty of all assumptions all the time. It is, however, important to recognise that the importance of the uncertainty is itself a value-laden question (Rudner, 1953).

Given the practical impossibility of including all plausible model choices and assumptions, the following approach, which starts from the risk of being wrong, might therefore represent a more sensible way forward for dealing with uncertainty and disagreement in IAM research:

(i) What outputs are likely to have a large impact/matter/are important?
(ii) What range of values of this/these outputs would have a significant impact (on action/the future)?
(iii) Are there any plausible sets of model choices and assumptions that could produce outputs in this range?
(iv) What is the (subjective or objective) likelihood of these model choices and assumptions vs. other (more commonly used) model choices and assumptions?
(v) What are the reasons why these model choices and assumptions are not currently used (epistemic, pragmatic, computational, social, cultural, etc.)?

While climate policy should be based on the best available knowledge, it is important to also ask whether the best available knowledge is good enough; that is, whether the knowledge provides a reliable source of information on which to base policies. A consideration of model dependencies in relation to plausible model choices and assumptions, and an assessment of sensitivities and uncertainties, this thesis has argued, represents a good starting point for beginning to answer this question.

Bibliography

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

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

Agnew, M., Schrattenholzer, L., & Voss, A. (1979a). *A model for energy supply systems alternatives and their general environmental impact* (WP-79-006). http://pure.iiasa.ac.at/1177/

Agnew, M., Schrattenholzer, L., & Voss, A. (1979b). *A Model for Energy Supply Systems Alternatives and Their General Environmental Impact* (WP-79-6; WP-79-6).

Allcott, H., & Greenstone, M. (2012). Is There an Energy Efficiency Gap ? *Journal of Economic Perspectives*, *26*(1), 3–28.

Altamirano, J., Drouet, L., Sceia, A., Thalmann, P., & Vielle, M. (2008). *Coupling GEMINI-E3 and MARKAL-CHRES to Simulate Swiss Climate Policies*.

Ambec, S., Cohen, M. A., Elgie, S., & Lanoie, P. (2013). The porter hypothesis at 20: Can environmental regulation enhance innovation and competitiveness? *Review of Environmental Economics and Policy*, *7*(1), 2–22. https://doi.org/10.1093/reep/res016

Anadon, L. D., Baker, E., & Bosetti, V. (2017). Integrating uncertainty into public energy research and development decisions. In *Nature Energy* (Vol. 2, Issue 5). https://doi.org/10.1038/nenergy.2017.71

Anderson, K. (2015). Duality in climate science. *Nature Geoscience*, *Dec*. https://doi.org/10.1038/nclimate2814

Anderson, K., & Jewell, J. (2019). Climate-policy models debated. *Nature*, *573*(19 September), 8–9.

Anderson, K., & Peters, G. (2016). The trouble with negative emissions. *Science*, *354*(6309), 182–183. https://doi.org/10.1126/science.aah4567

Arrow, K. (1962). The economic implications of learning by doing. *Review of Economic Studies*, *29*(3), 155–173. https://www.jstor.org/stable/pdf/2295952.pdf?refreqid=excelsior:bff9706f489bbcb7eef5ce3d67a6dd3a

Babaee, S., Nagpure, A. S., & Decarolis, J. (2014). How much do electric drive vehicles matter to future U.S. emissions? *Environmental Science and Technology*, *48*(3), 1382–1390. https://doi.org/10.1021/es4045677

Babiker, M., Gurgel, A., Paltsev, S., & Reilly, J. (2009). Forward-looking versus recursive-dynamic modeling in climate policy analysis: A comparison. *Economic Modelling*, *26*(6), 1341–1354. https://doi.org/10.1016/j.econmod.2009.06.009

Barker, T. (2004). The transition to sustainability: a comparison of General–Equilibrium and Space–Time–Economics approaches. In *Tyndall Centre for Climate Change Research* (Issue November). http://tyndall.ac.uk/sites/default/files/wp62.pdf

Barker, T., Anger, A., Chewpreecha, U., & Pollitt, H. (2012). A new economics approach to modelling policies to achieve global 2020 targets for climate stabilisation. *International Review of Applied Economics*, *26*(2), 205–221. https://doi.org/10.1080/02692171.2011.631901

Barker, T., Qureshi, M. S., & Köhler, J. (2006). The costs of greenhouse gas mitigation with induced technological change: a meta-analysis of estimates in the literature. In *Tyndall Centre Working Paper*.

Barker, T., & Şerban Scrieciu, S. (2010). Modeling low climate stabilization with E3MG: Towards a “new economics” approach to simulating energy-environment-economy system dynamics. *The Energy Journal*, *31*(SPECIAL ISSUE), 137–164. https://doi.org/10.5547/ISSN0195-6574-EJ-Vol31-NoSI-6

Baron, R., & Salles, J. M. (1991). Comparaison des travaux de prospective energetique along terme et evaluation de differents scenarios relatifs au probleme de l’effet de serre. *Energie Economie El Societe*, *91*(1).

Barron, R., & McJeon, H. (2015). The differential impact of low-carbon technologies on climate change mitigation cost under a range of socioeconomic and climate policy scenarios. *Energy Policy*, *80*, 264–274. https://doi.org/10.1016/j.enpol.2015.01.038

Bauer, N., Brecha, R. J., & Luderer, G. (2012). Economics of nuclear power and climate change mitigation policies. *Proceedings of the National Academy of Sciences*, *109*(42), 16805–16810. https://doi.org/10.1073/pnas.1201264109

Bauer, N., Mouratiadou, I., Luderer, G., Baumstark, L., Brecha, R. J., Edenhofer, O., & Kriegler, E. (2016). Global fossil energy markets and climate change mitigation – an analysis with REMIND. *Climatic Change*, *136*, 69–82. https://doi.org/10.1007/s10584-013-0901-6

Bauer, N., Rose, S. K., Fujimori, S., van Vuuren, D. P., Weyant, J., Wise, M., Cui, Y., Daioglou, V., Gidden, M. J., Kato, E., Kitous, A., Leblanc, F., Sands, R., Sano, F., Strefler, J., Tsutsui, J., Bibas, R., Fricko, O., Hasegawa, T., … Muratori, M. (2018). Global energy sector emission reductions and bioenergy use: overview of the bioenergy demand phase of the EMF-33 model comparison. *Climatic Change*. https://doi.org/10.1007/s10584-018-2226-y

Beaumais, O., & Zagame, P. (1993). 3.2 The Use of Economic Models for Analysing Environmental Problems: the Case of the Greenhouse Effect for France. In *The European Carbon Tax: An Economic Assessment*. https://doi.org/10.1007/978-94-011-1904-7

Beck, M., & Krueger, T. (2016). The epistemic, ethical, and political dimensions of uncertainty in integrated assessment modeling. *Wiley Interdisciplinary Reviews: Climate Change*, *7*(5), 627–645. https://doi.org/10.1002/wcc.415

Bibas, R., & Méjean, A. (2014). Potential and limitations of bioenergy for low carbon transitions. *Climatic Change*, *123*(3–4), 731–761. https://doi.org/10.1007/s10584-013-0962-6

Blanford, G. J., Kriegler, E., & Tavoni, M. (2014). Harmonization vs. fragmentation: Overview of climate policy scenarios in EMF27. *Climatic Change*, *123*(3–4), 383–396. https://doi.org/10.1007/s10584-013-0951-9

Blanford, G., Merrick, J., Richels, R., & Rose, S. (2014). Trade-offs between mitigation costs and temperature change. *Climatic Change*, *123*(3–4), 527–541. https://doi.org/10.1007/s10584-013-0869-2

Bohringer, C., & Loschel, A. (2006). Computable general equilibrium models for sustainability impact assessment: Status quo and prospects. *Ecological Economics*, *60*(1), 49–64. https://doi.org/10.1016/j.ecolecon.2006.03.006

Bosetti, V., Carraro, C., & Galeotti, M. (2006). The dynamics of carbon and energy intensity in a model of endogenous technical change. *The Energy Journal*, *27*(Special Issue: Endogenous Technological Change and the Economics of Atmospheric Stabilization), 191–206. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Dynamics+of+Carbon+and+Energy+Intensity+in+a+Model+of+Endogenous+Technical+Change#0

Bosetti, V., Carraro, C., Galeotti, M., Massetti, E., & Tavoni, M. (2006). WITCH A World Induced Technical Change Hybrid Model Emanuele Massetti. *The Energy Journal*, *27*, 13–37. https://doi.org/10.2139/ssrn.948382

Bosetti, V., Carraro, C., Massetti, E., & Tavoni, M. (Eds.). (2014). *Climate Change Mitigation, Technological Innovation and Adaptation: A new perspective on climate policy*. Edward Elgar.

Bosetti, V., Marangoni, G., Borgonovo, E., Diaz Anadon, L., Barron, R., McJeon, H. C., Politis, S., & Friley, P. (2015). Sensitivity to energy technology costs: A multi-model comparison analysis. *Energy Policy*, *80*, 244–263. https://doi.org/10.1016/j.enpol.2014.12.012

Bosquet, B. (2000). Environmental tax reform : does it work ? A survey of the empirical evidence. *Ecological Economics*, *34*, 19–32.

Boulanger, P.-M., & Bréchet, T. (2005). Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics*, *55*(3), 337–350. https://doi.org/10.1016/j.ecolecon.2005.07.033

Bouwman, A. F., Kram, T., Goldewijk, K. K., & Klein Goldewijk, K. (2006). Integrated modelling of global environmental change. An overview of IMAGE 2.4. In *Bilthoven The Netherlands Netherlands Environmental Assessment Agency MNP Retrieved June* (Vol. 15). www.mnp.nl/en

Bowen, A., & Fankhauser, S. (2011). The green growth narrative: Paradigm shift or just spin? *Global Environmental Change*, *21*(4), 1157–1159. https://doi.org/10.1016/j.gloenvcha.2011.07.007

Brenkert, A. L., Sands, R. D., Kim, S. H., & Pitcher, H. M. (2004). *Model Documentation: The Second Generation Model* (Issue October). http://www.globalchange.umd.edu/data/models/SGM-

Model_Documentation.pdf
Brenkert, A. L., Smith, S. J., Kim, S. H., & Pitcher, H. M. (2003). *Model Documentation for the MiniCAM*. http://www.globalchange.umd.edu/data/publications/MiniCAM.pdf
Butler, M. P., Reed, P. M., Fisher-Vanden, K., Keller, K., & Wagener, T. (2014). Identifying parametric controls and dependencies in integrated assessment models using global sensitivity analysis. *Environmental Modelling and Software*, *59*, 10–29. https://doi.org/10.1016/j.envsoft.2014.05.001
Bye, B., Kverndokk, S., & Rosendahl, K. E. (2002). Mitigation costs, distributional effects, and ancillary benefits of carbon policies in the Nordic countries, the U.K., and Ireland. *Mitigation and Adaptation Strategies for Global Change*, *7*(4), 339–366. https://doi.org/10.1023/A:1024741018194
Calvin, K., Clarke, L., Krey, V., Blanford, G., Jiang, K., Kainuma, M., Kriegler, E., Luderer, G., & Shukla, P. R. (2012). The role of Asia in mitigating climate change: Results from the Asia modeling exercise. *Energy Economics*, *34*(SUPPL. 3), S251–S260. https://doi.org/10.1016/j.eneco.2012.09.003
Calvin, K., Wise, M., Luckow, P., Kyle, P., Clarke, L., & Edmonds, J. (2016). Implications of uncertain future fossil energy resources on bioenergy use and terrestrial carbon emissions. *Climatic Change*, *136*(1), 57–68. https://doi.org/10.1007/s10584-013-0923-0
Cambridge Econometrics. (2019). *E3ME: Our Global Macro-econometric Model*. https://www.camecon.com/how/e3me-model/
Campolongo, F., & Braddock, R. (1999). Sensitivity analysis of the IMAGE greenhouse model. *Environmental Modelling and Software*, *14*(4), 275–282. https://doi.org/10.1016/S1364-8152(98)00079-6
Chan, G., & Anadon, L. D. (2016). *Improving Decision Making for Public R & D Investment in Energy : Utilizing Expert Elicitation in Parametric Models Gabriel Chan and Laura Diaz Anadon* (No. 1631; EPRG Working Paper, Issue December).
Chang, H.-J. (2014). *Economics: The User's Guide* (1st ed.). Bloomsbury Publishing USA.
Chappin, E. J. L., & Ligtvoet, A. (2014). Transition and transformation: A bibliometric analysis of two scientific networks researching socio-technical change. *Renewable and Sustainable Energy Reviews*, *30*, 715–723. https://doi.org/10.1016/j.rser.2013.11.013
Chen, W., Yin, X., & Zhang, H. (2016). Towards low carbon development in China: a comparison of national and global models. *Climatic Change*, *136*, 95–108. https://doi.org/10.1007/s10584-013-0937-7
Clarke, L., Edmonds, J., Krey, V., Richels, R., Rose, S., & Tavoni, M. (2009). International climate policy architectures: Overview of the EMF 22 International Scenarios. *Energy Economics*, *31*(SUPPL. 2), S64–S81. https://doi.org/10.1016/j.eneco.2009.10.013
Corbera, E., Calvet-Mir, L., Hughes, H., & Paterson, M. (2016). Patterns of authorship in the IPCC Working Group III report. *Nature Climate Change*, *6*(1), 94–99. https://doi.org/10.1038/nclimate2782
Crassous, R., Hourcade, J., & Sassi, O. (2006). Endogenous Structural Change and Climate Targets Modeling Experiments with Im ... *The Energy Journal*, *27*(Special Issue: Endogenous Technological Change and the Economics of Atmospheric Stabilization), 259–276. https://doi.org/10.5547/ISSN0195-6574-EJ-VolSI2006-NoSI1-13
De Cian, E., Sferra, F., & Tavoni, M. (2016). The influence of economic growth, population, and fossil fuel scarcity on energy investments. *Climatic Change*, *136*(1), 39–55. https://doi.org/10.1007/s10584-013-0902-5
DeCanio, S. J. (1997). ECONOMIC MODELING AND THE FALSE TRADEOFF BETWEEN ENVIRONMENTAL PROTECTION AND ECONOMIC GROWTH. *Contemporary Economic Policy*, *XV*(October). https://doi.org/10.1111/j.1465-7287.1997.tb00485.x
DeCarolis, J. (2011). Using modeling to generate alternatives (MGA) to expand our thinking on energy futures. *Energy Economics*, *33*(2), 145–152. https://doi.org/10.1016/j.eneco.2010.05.002
DeCarolis, J., Daly, H., Dodds, P., Keppo, I., Li, F., McDowall, W., Pye, S., Strachan, N., Trutnevyte, E., Usher, W., Winning, M., Yeh, S., & Zeyringer, M. (2017). Formalizing best practice for energy system optimization modelling. *Applied Energy*, *194*, 184–198. https://doi.org/10.1016/j.apenergy.2017.03.001
Decarolis, J., Hunter, K., & Sreepathi, S. (2012). The case for repeatable analysis with energy economy

optimization models. *Energy Economics*. https://doi.org/10.1016/j.eneco.2012.07.004
Dietz, S., Hepburn, C., & Stern, N. (2007). Economics, ethics and climate change. *Ethics and Climate Change*, 1–20. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1090572
Dodds, P. E., Keppo, I., & Strachan, N. (2015). Characterising the Evolution of Energy System Models Using Model Archaeology. *Environmental Modeling and Assessment*, *20*(2), 83–102. https://doi.org/10.1007/s10666-014-9417-3
Douglas, H. (2000). Inductive Risk and Values in Science. *Philosophy of Science*, *67*(4), 559–579.
Douglas, H. (2009). *Science, Policy and the Value-free Ideal*. Pittsburgh University Press.
Dowlatabadi, H. (1995). Integrated assessment models of climate change: An incomplete overview. *Energy Policy*, *23*(4), 289–296. http://www.sciencedirect.com/science/article/pii/030142159590155Z
Edenhofer, O., Hirth, L., Knopf, B., Pahle, M., Schlömer, S., Schmid, E., & Ueckerdt, F. (2013). On the economics of renewable energy sources. *Energy Economics*, *40*, S12–S23. https://doi.org/10.1016/j.eneco.2013.09.015
Edenhofer, O., Knopf, B., Leimbach, M., & Bauer, N. (2010). ADAM Modeling Comparison Project-Intentions and Prospects. *The Energy Journal*, *31*(Special Issue 1), 7–10. https://www.jstor.org/stable/pdf/41323489.pdf?refreqid=excelsior%3Ac4f5afcde0857c0d014b88d3c7dcb938
Edenhofer, O., Knopf, B., Leimbach, M., Bauer, N., Edenhofer, O., Knopf, B., Barker, T., Bellevrat, E., Château, B., Criqui, P., Isaac, M., Kitous, A., Kypreos, S., Leimbach, M., & Lessmann, K. (2010). Special Issue: The Economics of Low Stabilisation. In *The Energy Journal* (Vol. 31, Issue 1).
Edenhofer, O., Lessmann, K., Kemfert, C., Grubb, M., & Köhler, J. (2006). Induced technological chance: Exploring its implications for the economics of atmosferic stabilization. Synthesis report from the innovation modelling comparison project. *The Energy Journal*, *27*(Special Issue: Endogenous Technological Change and the Economics of Atmospheric Stabilization), 57–122.
Edmonds, J., & Reilly, J. (1983a). A long-term global energy- economic model of carbon dioxide release from fossil fuel use. *Energy Economics*, *5*(2), 74–88. https://doi.org/10.1016/0140-9883(83)90014-2
Edmonds, J., & Reilly, J. (1983b). Global Energy and CO2 to the Year 2050. *The Energy Journal*, *4*(3), 21–47. https://doi.org/10.5547/issn0195-6574-ej-vol4-no3-3
Edmonds, J., & Reilly, J. (1983c). Global energy production and use to the year 2050. *Energy*, *8*(6), 419–432. https://doi.org/10.1016/0360-5442(83)90064-6
Edwards, M. A., & Roy, S. (2017). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science*, *34*(1), 51–61. https://doi.org/10.1089/ees.2016.0223
Egli, F., Steffen, B., & Schmidt, T. S. (2018). A dynamic analysis of financing conditions for renewable energy technologies. *Nature Energy (Accepted), (accepted)*. https://doi.org/10.1038/s41560-018-0277-y
Elliott, K. C. (2017). *A Tapestry of Values: An Introduction to Values in Science*. Oxford University Press.
Enerdata. (2019). *POLES: Prospective Outlook on Long-term Energy Systems*. https://www.enerdata.net/solutions/poles-model.html
ETSAP. (2019). *MARKAL*. https://iea-etsap.org/index.php/etsap-tools/model-generators/markal
EU Science Hub. (2019). *GEM-E3. General Equilibirum Model for Economy-Energy-Environment*. https://ec.europa.eu/jrc/en/gem-e3/model
European Commission. (2018). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE, THE COMMITTEE OF THE REGIONS AND THE EUROPEAN INVESTMENT BANK. A Clean Planet for all. A European strategic long-t*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0773
Farmer, J. D., Hepburn, C., Mealy, P., & Teytelboym, A. (2015). A Third Wave in the Economics of Climate Change. *Environmental and Resource Economics*, *62*(2), 329–357. https://doi.org/10.1007/s10640-015-9965-2
Finon, D. (1974). Optimisation model for the French energy sector. *Energy Policy*, *2*(2), 136–151.

https://doi.org/10.1016/0301-4215(74)90005-6
Fragkos, P., Kouvaritakis, N., & Capros, P. (2015). Incorporating Uncertainty into World Energy Modelling: the PROMETHEUS Model. *Environmental Modeling and Assessment*, *20*(5), 549–569. https://doi.org/10.1007/s10666-015-9442-x
Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Quéré, C. Le, Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., … Zaehle, S. (2019). *Global Carbon Budget 2019*. *October*, 1–4.
Fujii, Y., & Yamaji, K. (1998). Assessment of technological options in the global energy system for limiting the atmospheric CO2 concentration. *Environmental Economics and Policy Studies*, *1*(2), 113–139. https://doi.org/10.1007/bf03353897
Funtowicz, S. O., & Ravetz, J. . (1990). *Uncertainty and Quality in Sicence for Policy*.
Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, *25*(7), 739–755. https://doi.org/10.1016/0016-3287(93)90022-L
Fuss, S., Canadell, J. G., Peters, G. P., Tavoni, M., Andrew, R. M., Ciais, P., Jackson, R. B., Jones, C. D., Kraxner, F., Nakicenovic, N., Le Quéré, C., Raupach, M. R., Sharifi, A., Smith, P., & Yamagata, Y. (2014). Betting on negative emissions. *Nature Climate Change*, *4*(Oct). https://doi.org/10.1038/nclimate2392
Gambhir, A., Butnar, I., Li, P. H., Smith, P., & Strachan, N. (2019). A review of criticisms of integrated assessment models and proposed approaches to address these, through the lens of BECCs. *Energies*, *12*(9), 1–21. https://doi.org/10.3390/en12091747
GCAM v5.1 Documentation. (2019). *GCAM Model Overview*. http://jgcri.github.io/gcam-doc/overview.html
Gerarden, T. D., Newell, R. G., & Stavins, R. N. (2017). Assessing the Energy-Efficiency Gap. *Journal of Economic Literature*, *55*(4), 1486–1525.
Glanemann, N., Willner, S. N., & Levermann, A. (2020). Paris Climate Agreement passes the cost-benefit test. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-019-13961-1
Goulder, L. H. (1995). Environmental Taxation and the Double Dividend. *International Tax and Public Finance*, *2*, 157–183. https://doi.org/10.1108/S0573-8555(2000)246
Grant, N., Hawkes, A., Napp, T., & Gambhir, A. (2020). The appropriate use of reference scenarios in mitigation analysis. *Nature Climate Change*, *10*(July), 1–6. https://doi.org/10.1038/s41558-020-0826-9
Grubb, M., Carraro, C., & Schellnhuber, J. (2006). Technological Change for Atmospheric Stabilization: Introductory Overview to the Innovation Modeling Comparison Project. *The Energy Journal*, *27*(Special Issue: Endogenous Technological Change and the Economics of Atmospheric Stabilization), 1–17. https://doi.org/10.5547/ISSN0195-6574-EJ-VolSI2006-NoSI1-1
Grubb, M., Köhler, J., & Anderson, D. (2002). Induced Technical Change in Energy and Environmental Modeling: Analytic Approaches and Policy Implications. *Annual Review of Energy and the Environment*, *27*(1), 271–308. https://doi.org/10.1146/annurev.energy.27.122001.083408
Grübler, A., Nakićenović, N., & Victor, D. G. (1999). Dynamics of energy technologies and global change. *Energy Policy*, *27*(5), 247–280. https://doi.org/10.1016/S0301-4215(98)00067-6
Guivarch, C., Crassous, R., & Sassi, O. (2011). The costs of climate policies in a second-best world with labour market imperfections. *Climate Policy*, *11*, 768–788. https://doi.org/10.3763/cpol.2009.0012
Häfele, W., & Manne, A. S. (1974). *Strategies for a transition from fossil to nuclear fuels*. https://doi.org/10.1016/0301-4215(75)90050-6
Hallegatte, S., Heal, G., Fay, M., & Treguer, D. (2012). *From Growth to Green Growth - A Framework* (Working Paper 17841). https://doi.org/10.1017/CBO9781107415324.004
Hausfather, Z., & Peters, G. P. (2020). Emissions – the 'business as usual' story is misleading. *Nature*, *577*(January), 618–620.
Hedenus, F., Johansson, D., & Lindgren, K. (2013). A Critical Assessment of Energy-economy-climate Models for Policy Analysis. *Journal of Applied Economics and Business Research*, *3*(2), 118–132. https://doi.org/10.1023/A:1019002620369
Hodges, J., & Dewar, J. (1992). *Is it you or your model talking?: A framework for model validation*. Rand Corporation. http://www.rand.org/content/dam/rand/pubs/reports/2006/R4114.pdf

Hope, C., Anderson, J., & Wenman, P. (2003). Policy analysis of the greenhouse effect. *Energy Policy*, *21*(3), 327–338. https://doi.org/10.1016/0301-4215(93)90253-c

Hourcade, J. C. (1993). Modelling long-run scenarios. Methodology lessons from a prospective study on a low CO2 intensive country. *Energy Policy*, *21*(3), 309–326. https://doi.org/10.1016/0301-4215(93)90252-B

Hu, M. C., & Hobbs, B. F. (2010). Analysis of multi-pollutant policies for the U.S. power sector under technology and policy uncertainty using MARKAL. *Energy*, *35*(12), 5430–5442. https://doi.org/10.1016/j.energy.2010.07.001

IAM helpful or not? (2015). [Editorial]. *Nature Climate Change*, *5*(2), 81–81. https://doi.org/10.1038/nclimate2526

IAMC. (2014). *AR5 Scenario Database*. https://secure.iiasa.ac.at/web-apps/ene/AR5DB/

IEA. (2019). *Global Energy & CO2 Status Report*. https://www.iea.org/geco/

IEA, & NEA. (2015). *Projected Costs of Generating Electricity 2015*. https://doi.org/10.1787/cost_electricity-2015-en

Iman, R. L., & Conover, W. J. (1982). A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables. *Communications in Statistics - Simulation and Computation*, *11*(3), 311–334. https://doi.org/10.1080/03610918208812265

Interagency Working Group on Social Cost of Carbon. (2010). *Technical Support Document : -- Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866* - (Issue February). http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Technical+Support+Document+:+-+Social+Cost+of+Carbon+for+Regulatory+Impact+Analysis#1

IPCC. (1990). *Climate Change: The IPCC Response Strategies*.

IPCC. (1992). *Climate Change 1992. The Supplementary Report to the IPCC Scientific Assessment* (J. T. Houghton, B. A. Callander, & S. K. Varney (Eds.)).

IPCC. (1995). *Climate Change 1995: Economic and Social Dimensions. Contribution of Working Group III to the Second Assessment Report on the Intergovernmental Panel on Climate Change.* https://doi.org/10.1016/S0959-3780(97)82915-9

IPCC. (2007). *Climate Change 2007: Mitigation of climate change. Working Group III Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* (B. Metz, O. Davidson, P. Bosch, R. Dave, & L. Meyer (Eds.)). Cambridge University Press.

IPCC. (2014a). *Climate Change 2014: Mitigation of Climate Change. Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.*

IPCC. (2014b). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (R. K. Pachauri; & L. Meyer (Eds.)). https://doi.org/10.1021/om00044a023

IPCC. (2018a). *Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change*, (V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, & T. Waterfield (Eds.)).

IPCC. (2018b). Summary for Policymakers. In V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, & T. Waterfield (Eds.), *Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change,*. https://www.ipcc.ch/site/assets/uploads/sites/2/2018/07/SR15_SPM_version_stand_alone_LR.pdf

Iyer, G., & Edmonds, J. (2018). Interpreting energy scenarios. *Nature Energy*, *3*(5), 357–358. https://doi.org/10.1038/s41560-018-0145-9

Iyer, G., Hultman, N., Fetter, S., & Kim, S. H. (2014). Implications of small modular reactors for climate change mitigation. *Energy Economics*, *45*, 144–154. https://doi.org/10.1016/j.eneco.2014.06.023

Jackson, R. B., Le Quéré, C., Andrew, R. M., Canadell, J. G., Korsbakken, J. I., Liu, Z., Peters, G. P., & Zheng, B. (2018). Global energy growth is outpacing decarbonization. *Environmental Research*

Letters, *13*(12). https://doi.org/10.1088/1748-9326/aaf303
Jacobs, M. (2013). Green growth. In J. Zsyman & M. Huberty (Eds.), *Handbook of Global Climate and Environmental Policy* (pp. 197–214). https://doi.org/10.1007/s10272-012-0415-x
Jaeger, C. C., Leonidas Paroussos, Diana Mangalagiu, Roland Kupers, Antoine Mandel, Tàbara, J. D., Jaeger, C. C., Paroussos, L., Mangalagiu, D., Kupers, R., & Mandel, A. (2011). *A New Growth Path for Europe*. http://www.euractiv.com/sites/all/euractiv/files/m30_Synthesis_0.pdf%5Cnhttp://edoc.gfz-potsdam.de/pik/get/4813/0/2bbb2541f5e08c473dd1a087176b3b49/4813.pdf
Jaeger, C. C., Schütze, F., Mielke, J., Steudle, G., Wolf, S., Fürst, S., Mangalagiu, D., Meißner, F., Paroussos, L., Schneider, C., Conci, S., Fragkiadakis, K., Nikiforakis, N., & Bibas, R. (2015). *Investment-oriented climate policy: An opportunity for Europe. Analysis of growth and employment effects of an EU climate target of 50% in 2030* (Issue September).
Jaffe, A. B., & Stavins, R. N. (1994). The energy-efficiency gap - What does it mean ? *Energy Policy*, *199422*(10), 804–810. https://scholar.harvard.edu/files/stavins/files/the_energy_efficiency_gap.pdf
Johnson, T., DeCarolis, J., Shay, C. L., Loughlin, D. H., Gage, C. L., & Vijay, S. (2006). *MARKAL scenario analyses of technology options for the electric sector: The impact on air quality* (EPA/600/R-06/114).
Jones, V. (2009). *The Green Collar Economy*. HarperCollins Publishers Inc.
Jun, M., Knutti, R., & Nychka, D. W. (2008). Spatial analysis to quantify numerical model bias and dependence: How many climate models are there? *Journal of the American Statistical Association*, *103*(483), 934–947. https://doi.org/10.1198/016214507000001265
Kainuma, M., Matsuoka, Y., & Morita, T. (2003). *Climate Policy Assessment*.
Kainuma, M., Matsuoka, Y., & Morita, T. yuki. (1995). End-use Energy Model for Analysing the Policy Options to Reduce Greenhouse Gas Emissions. In *IIASA Working Paper* (WP-95-121; IIASA Working Paper).
Kann, a, & Weyant, J. P. (2000). Approaches for performing uncertainty analysis in large-scale energy/economic policy models. *Environmental Modeling & Assessment*, *5*(1), 29–46. https://doi.org/10.1023/A:1019041023520
Keepin, B., & Wynne, B. (1984). Technical analysis of IIASA energy scenarios. *Nature*, *312*(5996), 691–695. https://doi.org/10.1038/312691a0
Keppo, I., & Strubegger, M. (2010). Short term decisions for long term problems - The effect of foresight on model based energy systems analysis. *Energy*, *35*(5), 2033–2042. https://doi.org/10.1016/j.energy.2010.01.019
Kincaid, H., Dupré, J., & Wylie, A. (Eds.). (2007). *Value-Free Science? Ideals and Illusions*. Oxford University Press.
Knutti, R. (2018). A wider role for climate scenarios. *Nature Sustainability*, *1*(5), 214–215. https://doi.org/10.1038/s41893-018-0067-3
Knutti, R., Masson, D., & Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, *40*(6), 1194–1199. https://doi.org/10.1002/grl.50256
Krey, V., Guo, F., Kolp, P., Zhou, W., Schaeffer, R., Awasthy, A., Bertram, C., de Boer, H. S., Fragkos, P., Fujimori, S., He, C., Iyer, G., Keramidas, K., Köberle, A. C., Oshiro, K., Reis, L. A., Shoai-Tehrani, B., Vishwanathan, S., Capros, P., … van Vuuren, D. P. (2019a). Looking under the hood: A comparison of techno-economic assumptions across national and global integrated assessment models. Appendix A. *Energy*, 1254–1267. https://doi.org/10.1016/j.energy.2018.12.131
Krey, V., Guo, F., Kolp, P., Zhou, W., Schaeffer, R., Awasthy, A., Bertram, C., de Boer, H. S., Fragkos, P., Fujimori, S., He, C., Iyer, G., Keramidas, K., Köberle, A. C., Oshiro, K., Reis, L. A., Shoai-Tehrani, B., Vishwanathan, S., Capros, P., … van Vuuren, D. P. (2019b). Looking under the hood: A comparison of techno-economic assumptions across national and global integrated assessment models. *Energy*, *172*, 1254–1267. https://doi.org/10.1016/j.energy.2018.12.131
Krey, V., Luderer, G., Clarke, L., & Kriegler, E. (2014). Getting from here to there - energy technology transformation pathways in the EMF27 scenarios. *Climatic Change*, *123*(3–4), 369–382. https://doi.org/10.1007/s10584-013-0947-5
Krey, V., & Riahi, K. (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*(SUPPL. 2), S94–S106. https://doi.org/10.1016/j.eneco.2009.07.001

Kriegler, E., Riahi, K., Bauer, N., Schwanitz, V. J., Petermann, N., Bosetti, V., Marcucci, A., Otto, S., Paroussos, L., Rao, S., Currás, T. A., Ashina, S., Bollen, J., Eom, J., Hamdi-Cherif, M., Longden, T., Kitous, A., Méjean, A., Sano, F., … Edenhofer, O. (2015a). Making or breaking climate targets: The AMPERE study on staged accession scenarios for climate policy. Supplementary Online Material. *Technological Forecasting and Social Change*, *90*, 322–326. https://doi.org/10.1016/j.techfore.2015.01.019

Kriegler, E., Riahi, K., Bauer, N., Schwanitz, V. J., Petermann, N., Bosetti, V., Marcucci, A., Otto, S., Paroussos, L., Rao, S., Currás, T. A., Ashina, S., Bollen, J., Eom, J., Hamdi-Cherif, M., Longden, T., Kitous, A., Méjean, A., Sano, F., … Edenhofer, O. (2015b). Making or breaking climate targets: The AMPERE study on staged accession scenarios for climate policy. *Technological Forecasting & Social Change*, *90*, 24–44. https://doi.org/10.1016/j.techfore

Kriegler, E., Tavoni, M., Aboumahboub, T., Luderer, G., Calvin, K., Demaere, G., Krey, V., Riahi, K., Rösler, H., Schaeffer, M., & Van Vuuren, D. P. (2013). What does the 2°c target imply for a global climate agreement in 2020? the limits study on durban platform scenarios. *Climate Change Economics*, *04*(04), 1340008. https://doi.org/10.1142/S2010007813400083

Kriegler, E., Weyant, J. P., Blanford, G. J., Krey, V., Clarke, L., Edmonds, J., Fawcett, A., Luderer, G., Riahi, K., Richels, R., Rose, S. K., Tavoni, M., & van Vuuren, D. P. (2014). The role of technology for achieving climate policy objectives: Overview of the EMF 27 study on global technology and climate policy strategies. *Climatic Change*, *123*(3–4), 353–367. https://doi.org/10.1007/s10584-013-0953-7

Krugman, P. (2014). Paradigming Is Hard. *Paul Krugman Blog*, *April 28*. http://krugman.blogs.nytimes.com/2014/04/28/paradigming-is-hard/

Kuhn, T. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Kypreos, S., & Bahn, O. (2003). A MERGE model with endogenous technological progress. *Environmental Modeling and Assessment*, *8*(3), 249–259. https://doi.org/10.1023/A:1025551408939

Labriet, M., Kanudia, A., & Loulou, R. (2012). Climate mitigation under an uncertain technology future: A TIAM-World analysis. *Energy Economics*, *34*(SUPPL. 3), S366–S377. https://doi.org/10.1016/j.eneco.2012.02.016

Larkin, A., Kuriakose, J., Sharmina, M., & Anderson, K. (2018). What if negative emission technologies fail at scale? Implications of the Paris Agreement for big emitting nations. *Climate Policy*, *18*(6), 690–714. https://doi.org/10.1080/14693062.2017.1346498

Lehtveer, M., & Hedenus, F. (2015). How much can nuclear power reduce climate mitigation cost? - Critical parameters and sensitivity. *Energy Strategy Reviews*, *6*, 12–19. https://doi.org/10.1016/j.esr.2014.11.003

Leimbach, M., Bauer, N., Baumstark, L., & Edenhofer, O. (2010). Mitigation costs in a globalized world: Climate policy analysis with REMIND-R. *Environmental Modeling and Assessment*, *15*(3), 155–173. https://doi.org/10.1007/s10666-009-9204-8

Lemoine, D., & McJeon, H. C. (2013). Trapped between two tails: Trading off scientific uncertainties via climate targets. *Environmental Research Letters*, *8*(3). https://doi.org/10.1088/1748-9326/8/3/034019

Lesourd, J.-B., Percebois, J., & Valette, F. (1996). *Models for Energy Policy*. https://doi.org/10.4324/9780203982525

Löschel, A. (2002). Technological change in economic models of environmental policy: A survey. *Ecological Economics*, *43*(2–3), 105–126. https://doi.org/10.1016/S0921-8009(02)00209-4

Loucks, D. P., van Beek, E., Stedinger, J. R., Dijkman, J. P. M., & Villars, M. T. (2005). Model Sensitivity and Uncertainty Analysis. In *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications* (pp. 254–290). https://doi.org/ISBN: 92-3-103998-9

Loulou, R, Goldstein, G., Kanudia, A., Lettila, A., & Remme, U. (2016). *Documentation for the TIMES model: PART I* (Issue July). https://iea-etsap.org/docs/Documentation_for_the_TIMES_Model-Part-I_July-2016.pdf

Loulou, Richard, Goldstein, G., & Noble, K. (2004). *Documentation for the MARKAL Family of Models Part II: MARKAL-MACRO*. *October*, 1–386. http://www.iea-etsap.org/web/MrklDoc-I_StdMARKAL.pdf
Loulou, Richard, & Labriet, M. (2008). ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure. *Computational Management Science*, *5*(1–2), 7–40. https://doi.org/10.1007/s10287-007-0046-z
Lübbe, H. (1986). Scientific practice and responsibility. In M. C. Doser & J. N. Kray (Eds.), *Facts and values* (pp. 81–95). Dortrecht: Martinus Nijhoff.
Lucas, P. L., Shukla, P. R., Chen, W., van Ruijven, B. J., Dhar, S., den Elzen, M. G. J., & van Vuuren, D. P. (2013). Implications of the international reduction pledges on long-term energy system changes and costs in China and India. *Energy Policy*, *63*, 1032–1041. https://doi.org/10.1016/j.enpol.2013.09.026
Luderer, G., Bertram, C., Calvin, K., De Cian, E., & Kriegler, E. (2016). Implications of weak near-term climate policies on long-term mitigation pathways. *Climatic Change*, *136*(1), 127–140. https://doi.org/10.1007/s10584-013-0899-9
Luderer, G., Bosetti, V., Jakob, M., Leimbach, M., Steckel, J. C., Waisman, H., & Edenhofer, O. (2012). The economics of decarbonizing the energy system-results and insights from the RECIPE model intercomparison. *Climatic Change*, *114*(1), 9–37. https://doi.org/10.1007/s10584-011-0105-x
Luderer, G., Leimbach, M., Bauer, N., & Kriegler, E. (2011). *Description of the ReMIND-R model.*
Luderer, G., Leimbach, M., Bauer, N., Kriegler, E., Aboumahboub, T., Curras, T. A., Baumstark, L., Bertram, C., Giannousakis, A., Hilaire, J., Klein, D., Mouratiadou, I., Pietzcker, R., Piontek, F., Roming, N., Schultes, A., Schwanitz, V. J., & Strefler, J. (2013). *Description of the ReMIND model (Version 1.5)* (pp. 1–33).
Lurz, J., Smith, S. J., Edmonds, J., Kim, S. H., & Wise, M. (2006). The objECTS Framework for integrated Assessment: Hybrid Modeling of Transportation. *The Energy Journal*, *27*(Special Issue: Hybrid Modeling of Energy-Environment Policies: Reconciling Bottom-up and Top-down), 63–91. https://doi.org/10.5547/issn0195-6574-ej-volsi2006-nosi2-4
Manne, A. S., Mendelsohn, R., & Richels, R. (1995). MERGE: A model for evaluating regional and global effects of GHG Reduction Policies. *Energy Policy*, *23*(I), 17–34.
Manne, A. S., & Richels, R. G. (1990). The Costs of Reducing U.S. CO2 Emissions - Further Sensitivity Analyses. *The Energy Journal*, *11*(4), 69–78. https://doi.org/10.1007/978-1-4757-2161-4_5
Manne, A. S., & Richels, R. G. (1992). *Buying greenhouse insurance: the economic costs of carbon dioxide emission limits*. MIT Press.
Manne, A. S., Richels, R. G., & Hogan, W. W. (1990). CO2 Emission Limits: An Economic Cost Analysis for the USA. *The Energy Journal*, *11*(2), 51. https://doi.org/10.2307/41322654
Manne, A. S., & Wene, C.-O. (1992). *MARKAL-MACRO: A linked model for energy-economy analysis*. https://doi.org/10.1081/E-ENN3-120050114
Marchetti, C., & Nakicenovic, N. (1979). The Dynamics of Energy Systems and the Logistic Substitution Model. In *IIASA Administrative Report*.
Maréchal, K. (2007). The economics of climate change and the change of climate in economics. *Energy Policy*, *35*(10), 5181–5194. https://doi.org/10.1016/j.enpol.2007.05.009
Marechal, K., & Lazaric, N. (2010). Overcoming inertia: Insights from evolutionary economics into improved energy and climate policies. *Climate Policy*, *10*(1), 103–119. https://doi.org/10.3763/cpol.2008.0601
Mark, H., Huan, G., Mandell, J., & Zysman, J. (2011). Shaping the green growth economy. In *The Berkeley Roundtable on the International Economy*.
Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford University Press.
Masson, D., & Knutti, R. (2011). Climate model genealogy. *Geophysical Research Letters*, *38*(8), 1–4. https://doi.org/10.1029/2011GL046864
Matsuoka, Y., Kainuma, M., & Morita, T. (1995). Scenario analysis of global warming using the Asian Pacific Integrated Model (AIM). *Energy Policy*, *23*(4–5), 357–371. https://doi.org/10.1016/0301-4215(95)90160-9
Mccollum, D. L., Gambhir, A., Rogelj, J., & Wilson, C. (2020). Energy modellers should explore extremes more systematically in scenarios. *Nature Energy*. https://doi.org/10.1038/s41560-020-

0555-3
Mcdowall, W., Trutnevyte, E., Tomei, J., & Keppo, I. (2014). *Reflecting on scenarios. UKERC energy systems theme working paper* (Issue June). http://discovery.ucl.ac.uk/1457412/1/McDowall_Reflecting on Scenarios Working Paper-3.pdf
McJeon, H. C., Lempert, R. J., Clarke, L., Bryant, B. P., Hackbarth, A., Kyle, P., & Wise, M. (2010). Technology interactions among low-carbon energy technologies: What can we learn from a large number of scenarios? *Energy Economics*, *33*(4), 619–631. https://doi.org/10.1016/j.eneco.2010.10.007
Mercure, J. F. (2011). Global electricity technology substitution model with induced technological change. *Tyndall Working Paper*, *148*.
Mercure, J. F. (2012). FTT:Power : A global model of the power sector with induced technological change and natural resource depletion. *Energy Policy*, *48*, 799–811. https://doi.org/10.1016/j.enpol.2012.06.025
Mercure, J. F. (2015). An age structured demographic theory of technological change. *Journal of Evolutionary Economics*, *25*, 787–820.
Mercure, J. F., Knobloch, F., Pollitt, H., Lewney, R., Rademakers, K., Eichler, L., van der Laan, J., & Paroussos, L. (2016). *Policy-induced energy technological innovation and finance for low-carbon economic growth*.
Mercure, J. F., Pollitt, H., Bassi, A. M., Viñuales, J. E., & Edwards, N. R. (2016). Modelling complex systems of heterogeneous agents to better design sustainability transitions policy. *Global Environmental Change*, *37*, 102–115. https://doi.org/10.1016/j.gloenvcha.2016.02.003
Mercure, J. F., Pollitt, H., Chewpreecha, U., Salas, P., Foley, A. M., Holden, P. B., & Edwards, N. R. (2014). The dynamics of technology diffusion and the impacts of climate policy instruments in the decarbonisation of the global electricity sector. *Energy Policy*, *73*, 686–700. https://doi.org/10.1016/j.enpol.2014.06.029
Mercure, J. F., Pollitt, H., Edwards, N. R., Holden, P. B., Chewpreecha, U., Salas, P., Lam, A., Knobloch, F., & Vinuales, J. E. (2018). Environmental impact assessment for climate change policy with the simulation-based integrated assessment model E3ME-FTT-GENIE. *Energy Strategy Reviews*, *20*, 195–208. https://doi.org/10.1016/j.esr.2018.03.003
Mercure, J. F., Pollitt, H., Viñuales, J. E., Edwards, N. R., Holden, P. B., Chewpreecha, U., Salas, P., Sognnaes, I., Lam, A., & Knobloch, F. (2018). Macroeconomic impact of stranded fossil fuel assets. *Nature Climate Change*, *8*(7), 588–593. https://doi.org/10.1038/s41558-018-0182-1
Mercure, J. F., & Salas, P. (2012). An assessement of global energy resource economic potentials. *Energy*, *46*(1), 322–336. https://doi.org/10.1016/j.energy.2012.08.018
Messner, S., & Schrattenholzer, L. (2000). MESSAGE–MACRO: linking an energy supply model with a macroeconomic module and solving it iteratively Sabine. *Energy*, *25*, 267–282. https://doi.org/10.1016/S0360-5442(99)00063-8
Morgan, M., & Henrion, M. (1990). *Uncertainty. A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press.
Morgenstern, O. (1963). *On the Accuracy of Economic Observations* (2nd ed.). Princeton University Press.
Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. https://doi.org/10.1080/00401706.1991.10484804
Munda, G. (1996). Cost-benefit analysis in integrated environmental assessment: some methodological issues. *Ecological Economics*, *19*(2), 157–168. https://doi.org/10.1016/0921-8009(96)00048-1
Mundaca, L., & Santi, F. (2004). *Quantitative assessment of selected policy instruments using the Western European MARKAL model. Phase III EU SAVE White and Green Project: Comparison of market-based instruments to promote energy efficiency* (No. 2004–10; IIIEE Reports).
Nakicenovic, N., & Swart, R. (2000). IPCC Special Report on Emissions Scenarios: A special report of Working Group III of the Intergovernmental Panel on Climate Change. In *Emissions Scenarios*.
Nemet, G. (2013). Technological Change and Climate Change Policy. In *Encyclopedia of Energy, Natural Resource and Environmental Economics* (pp. 97–112). Elsevier Inc. https://doi.org/10.4324/9781315736761
New Economics Foundation. (2008). *A Green New Deal: Joined-up policies to solve the triple crunch of the credit crisis, climate change and high oil prices*.

Nikas, A., Doukas, H., & Papandreou, A. (2019). A Detailed Overview and Consistent Classification of Climate-Economy Models. In *Understanding Risks and Uncertainties in Energy and Climate Policy*. https://doi.org/10.1007/978-3-030-03152-7

Nordhaus, W. (1991a). To Slow or Not to Slow: The Economics of The Greenhouse Effect. *The Economic Journal*, *101*(407), 920. https://doi.org/10.2307/2233864

Nordhaus, W. (1991b). To Slow or Not to Slow. *The Economic Journal*, *101*(407), 920–937. http://www.jstor.org/stable/2233864

Nordhaus, W. (1992). An Optimal Transition Path for Controlling Greenhouse Gases. *Science*, *258*(5086), 1315–1319. https://doi.org/10.1126/science.258.5086.1315

Nordhaus, W. (2007). A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, *XLV*(September), 686–702. http://www.jstor.org/stable/10.2307/27646843

Nordhaus, W. (2017). *NBER Reporter 2017 Number 3*. https://doi.org/10.1360/zd-2013-43-6-1064

OECD. (2011). *Towards Green Growth: Summary for policy makers*. https://doi.org/10.1787/9789264111318-en

Olaleye, O., & Baker, E. (2015). Large scale scenario analysis of future low carbon energy options. *Energy Economics*, *49*, 203–216. https://doi.org/10.1016/j.eneco.2015.02.006

Ondraczek, J., Komendantova, N., & Patt, A. (2015). WACC the dog: The effect of financing costs on the levelized cost of solar PV power. *Renewable Energy*, *75*, 888–898. https://doi.org/10.1016/j.renene.2014.10.053

Owens, S. (2015). 1 Knowledge, Advice, and Policy. In *Knowledge, Policy, and Expertise: The UK Royal Commission on Environmental Pollution 1970-2011*.

Oxford Dictionary of English (3rd ed.). (2010). Oxford University Press.

Parkinson, S., & Young, P. (1998). Uncertainty and sensitivity in global carbon cycle modelling. *Climate Research*, *9*(3), 157–174. https://doi.org/10.3354/cr009157

Peace, J., & Weyant, J. P. (2008). *Insights Not Numbers: The Appropriate Use of Economic Models* (White Paper – Pew Center on Global Climate Change).

Perez, C. (2009). Technological revolutions and techno-economic paradigms. *Cambridge Journal of Economics*, *34*(1), 185–202. https://doi.org/10.1093/cje/bep051

Perez, C. (2017). *Is Smart Green Growth the Solution? Lessons from History* (WP 2017-01; Beyond the Technological Revolution (BTTR)). http://beyondthetechrevolution.com/wp-content/uploads/2014/10/BTTR_WP_2017-1.pdf

Peters, M., Schmidt, T. S., Wiederkehr, D., & Schneider, M. (2011). Shedding light on solar technologies-A techno-economic assessment and its policy implications. *Energy Policy*, *39*(10), 6422–6439. https://doi.org/10.1016/j.enpol.2011.07.045

Pietzcker, R. C., Ueckerdt, F., Carrara, S., de Boer, H. S., Després, J., Fujimori, S., Johnson, N., Kitous, A., Scholz, Y., Sullivan, P., & Luderer, G. (2017). System integration of wind and solar power in integrated assessment models: A cross-model evaluation of new approaches. *Energy Economics*, *64*, 583–599. https://doi.org/10.1016/j.eneco.2016.11.018

Pindyck, R. S. (2013). Climate Change Policy: What Do the Models Tell Us? *Journal of Economic Literature*, *51*(3), 860–872. https://doi.org/10.1257/jel.51.3.860

Porter, M. E., & Van Der Linde, C. (1995). Toward a New Conception of the Environment-Competitiveness Relationship. *The Journal of Economic Perspectives*, *9*(4), 97–118.

Pratley, N. (2018, September). Putting workers on the board is a bitter but necessary pill. *The Guardian*. https://www.theguardian.com/business/nils-pratley-on-finance/2018/sep/05/putting-workers-on-the-board-is-a-bitter-but-necessary-pill

Prinn, R., Paltsev, S., Sokolov, A., Sarofim, M., Reilly, J., & Jacoby, H. (2011). Scenarios with MIT integrated global systems model: Significant global warming regardless of different approaches. *Climatic Change*, *104*(3–4), 515–537. https://doi.org/10.1007/s10584-009-9792-y

Pye, S., Sabio, N., & Strachan, N. (2015). An integrated systematic analysis of uncertainties in UK energy transition pathways. *Energy Policy*, *87*, 673–684. https://doi.org/10.1016/j.enpol.2014.12.031

Qu, W., Barney, G. O., Symalla, D., & Martin, L. (1998). The Threshold 21: National Sustainable Development Model. In *Integrated Global Models of Sustainable Development - Vol. II: Vol. II*.

Quéré, C. Le, Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., … 1Tyndall. (2018). Global Carbon Budget 2018. *Earth System Science Data*, *10*, 2141–2194. https://doi.org/10.5194/essd-8-605-2016

Repetto, R., & Austin, D. (1997). *The Costs of Climate Protection: A Guide for the Perplexed*.

Rezai, A., Taylor, L., & Mechler, R. (2013). Ecological macroeconomics: An application to climate change. *Ecological Economics*, *85*, 69–76. https://doi.org/10.1016/j.ecolecon.2012.10.008

Riahi, K., Kriegler, E., Johnson, N., Bertram, C., den Elzen, M., Eom, J., Schaeffer, M., Edmonds, J., Isaac, M., Krey, V., Longden, T., Luderer, G., Méjean, A., McCollum, D. L., Mima, S., Turton, H., van Vuuren, D. P., Wada, K., Bosetti, V., … Edenhofer, O. (2015a). Locked into Copenhagen pledges - Implications of short-term emission targets for the cost and feasibility of long-term climate goals. Supplementary Material. *Technological Forecasting and Social Change*, *90*(PA), 8–23. https://doi.org/10.1016/j.techfore.2013.09.016

Riahi, K., Kriegler, E., Johnson, N., Bertram, C., den Elzen, M., Eom, J., Schaeffer, M., Edmonds, J., Isaac, M., Krey, V., Longden, T., Luderer, G., Méjean, A., McCollum, D. L., Mima, S., Turton, H., van Vuuren, D. P., Wada, K., Bosetti, V., … Edenhofer, O. (2015b). Locked into Copenhagen pledges - Implications of short-term emission targets for the cost and feasibility of long-term climate goals. *Technological Forecasting and Social Change*, *90*(PA), 8–23. https://doi.org/10.1016/j.techfore.2013.09.016

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., & Rafaj, P. (2011). RCP 8.5-A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, *109*(1), 33–57. https://doi.org/10.1007/s10584-011-0149-y

Robinson, J. B. (1982). Energy backcasting. A proposed method of policy analysis. *Energy Policy*, *10*(4), 337–344. https://doi.org/10.1016/0301-4215(82)90048-9

Robinson, J. B. (1990). Futures Under Glass: A recipe for people who hate to predict. *Futures*, *October*. https://ac.els-cdn.com/001632879090018D/1-s2.0-001632879090018D-main.pdf?_tid=55e58e30-26e0-4b94-98b3-da67c4d762fb&acdnat=1538400194_f6ce99f7b69b177068d05071f9972f80

Rockström, J., Gaffney, O., Rogelj, J., Meinshausen, M., Nakicenovic, N., & Schellnhuber, H. J. (2017). A roadmap for rapid decarbonization. *Science*, *355*(6331), 1269–1271. https://doi.org/10.1126/science.aah3443

Rogelj, J., McCollum, D. L., Reisinger, A., Meinshausen, M., & Riahi, K. (2013). Probabilistic cost estimates for climate change mitigation. *Nature*, *493*(7430), 79–83. https://doi.org/10.1038/nature11787

Rogers, E. (1962). *Diffusion of Innovations* (1st ed.). Free Press of Glencoe.

Rosen, R. A. (2015). Critical review of: "Making or breaking climate targets — the AMPERE study on staged accession scenarios for climate policy." *Technological Forecasting & Social Change*, *96*, 322–326. https://doi.org/10.1016/j.techfore.2015.01.019

Rosen, R. A., & Guenther, E. (2015). The economics of mitigating climate change: What can we know? *Technological Forecasting and Social Change*, *91*, 93–106. https://doi.org/10.1016/j.techfore.2014.01.013

Rotmans, J. (1990). *Image : An Integrated Model to Assess the Greenhouse Effect*.

Rubin, E. S., Azevedo, I. M. L., Jaramillo, P., & Yeh, S. (2015). A review of learning rates for electricity supply technologies. *Energy Policy*, *86*, 198–218. https://doi.org/10.1016/j.enpol.2015.06.011

Rudner, R. (1953). The Scientist Qua Scientist Makes Value Judgments. *Philosophy of Science*, *20*(1), 1–6.

Salas, P. (2017). *The Effects of Uncertainty in the Technological Transitions of the Power Sector* (Issue March) [Cambridge]. https://www.repository.cam.ac.uk/bitstream/handle/1810/265885/Thesis_Pablo_Salas.pdf?sequence=1&isAllowed=y

Saltelli, A., Ratto, M., Tarantola, S., & Campolongo, F. (2006). Sensitivity analysis practices: Strategies for model-based inference. *Reliability Engineering and System Safety*, *91*(10–11), 1109–1125. https://doi.org/10.1016/j.ress.2005.11.014

Samadi, S. (2018). The experience curve theory and its application in the field of electricity generation technologies – A literature review. *Renewable and Sustainable Energy Reviews*, *82*(June 2016), 2346–2364. https://doi.org/10.1016/j.rser.2017.08.077

Sanstad, A. H., & Greening, L. A. (1998). Economic models for climate policy analysis: A critical discussion. *Environmental Modeling and Assessment*, *3*, 3–18. https://link.springer.com/content/pdf/10.1023%2FA%3A1019002620369.pdf

Sassi, O., Crassous, R., Hourcade, J. C., Gitz, V., Waisman, H., & Guivarch, C. (2010). IMACLIM-R: a modelling framework to simulate sustainable development pathways. *International Journal of Global Environmental Issues*, *10*(1/2), 5–24. https://doi.org/10.1504/ijgenvi.2010.030566

Schneider, S. H. (1997). Integrated assessment modeling of global climate change: Transparent rational tool for policy making or opaque screen hiding value-laden assumptions? *Environmental Modeling & Assessment*, *2*, 229–249. http://link.springer.com/article/10.1023/A:1019090117643

Schneider, S. H. (2001). What Is “ Dangerous ” Climate Change ? *Nature*, *411*, 17–19. http://stephenschneider.stanford.edu

Schumpeter, J. (1942). *Capitalism, Socialism and Democracy*. Harper.

Schumpeter, J. A. (1983). *The Theory of Economic Development: An inquiry into Profits, Capital, Credit, Interest, and the Business Cycle* (reprint, r). Transaction Publishers.

Scrieciu, S. S. (2007). The inherent dangers of using computable general equilibrium models as a single integrated modelling framework for sustainability impact assessment. A critical note on Böhringer and Löschel (2006). *Ecological Economics*, *60*(4), 678–684. https://doi.org/10.1016/j.ecolecon.2006.09.012

Scrieciu, S. S., Rezai, A., & Mechler, R. (2013). On the economic foundations of green growth discourses: The case of climate change mitigation and macroeconomic dynamics in economic modeling. *Wiley Interdisciplinary Reviews: Energy and Environment*, *2*(3), 251–268. https://doi.org/10.1002/wene.57

Shackley, S., Young, P., Parkinson, S., & Wynne, B. (1998). UNCERTAINTY, COMPLEXITY AND CONCEPTS OF GOOD SCIENCE IN CLIMATE CHANGE MODELLING: ARE GCMs THE BEST TOOLS? *Climatic Change*, *38*, 159–205.

Smith, P., Davis, S. J., Creutzig, F., Fuss, S., Minx, J., Gabrielle, B., Kato, E., Jackson, R. B., Cowie, A., Kriegler, E., Van Vuuren, D. P., Rogelj, J., Ciais, P., Milne, J., Canadell, J. G., McCollum, D., Peters, G., Andrew, R., Krey, V., … Yongsung, C. (2016). Biophysical and economic limits to negative CO2 emissions. *Nature Climate Change*, *6*(1), 42–50. https://doi.org/10.1038/nclimate2870

Spencer, T. A., Chancel, L., & Guerin, E. (2012). Exiting the Crisis in the Right Direction: A Sustainable and Shared Prosperity Plan for Europe. *SSRN Electronic Journal*, *April 2012*. https://doi.org/10.2139/ssrn.2058200

Stanford University. (2019). *MERGE Documentation*. https://web.stanford.edu/group/MERGE/biblio.htm

Stanton, E. A., Ackerman, F., & Kartha, S. (2008). Inside the integrated assessment models: Four issues in climate economics. In *Stockholm Environment Institute*. https://doi.org/10.3763/cdev.2009.0015

Stehfest, E., Vuuren, D. van, Kram, T., Bouwman, L., Alkemade, R., Bakkenes, M., Biemans, H., Bouwman, A., Elzen, M. den, Janse, J., Lucas, P., Minnen, J. van, Müller, C., & Prins, A. G. (2014). *Integrated Assessment of Global Environmental Change with Image 3.0: Model description and policy applications*. PBL Netherlands Environmental Assessment Agency. https://www.pbl.nl/sites/default/files/cms/publicaties/pbl-2014-integrated assessment of global environmental change with image30_735.pdf

Stern, N. (2009). *Time for a green industrial revolution*. New Scientist. https://institutions.newscientist.com/article/mg20126926-600-comment-time-for-a-green-industrial-revolution/

Stern, N. (2013a). *Ethics, Equity and the Economics of Climate Change Paper 1 : Science and Philosophy* (No. 97a; Centre for Climate Change Economics and Policy). http://www.lse.ac.uk/GranthamInstitute/publications/WorkingPapers/Papers/80-89/WP84a-Ethics,-equity-and-the-economics-of-climate-change---Paper-1-science-and-philosophy.pdf

Stern, N. (2013b). The Structure of Economic Modeling of the Potential Impacts of Climate Change: Grafting Gross Underestimation of Risk onto Already Narrow Science Models. *Journal of Economic Literature*, *51*(3), 838–859. https://doi.org/10.1257/jel.51.3.838

Stern, N. (2014). Chapter 5: Economics of Change. In *The New Climate Economy*.

Stern, N. (2016). Current climate models are grossly misleading. *Nature*, *530*, 407–409. https://doi.org/10.1038/530407a
Stirling, A. (2010). Keep it complex. *Nature*, *468*.
Tavoni, M., Kriegler, E., ABOUMAHBOUB, T., Calvin, K., Maere, G. De, Jewell, J., Kober, T., Lucas, P., Luderer, G., McCollum, D., Marangoni, G., Riahi, K., & Vuuren, D. van. (2014). *The distribution of the major economies' effort in the Durban platform scenarios*. http://www.internationalenergyworkshop.org/docs/IEW 2013_4D4Tavoni.pdf
Tebaldi, C., & Knutti, R. (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. https://doi.org/10.1098/rsta.2007.2076
Trutnevyte, E. (2016). Does cost optimization approximate the real-world energy transition? *Energy*, *106*, 182–193. https://doi.org/10.1016/j.energy.2016.03.038
Turner, A. (2012). *Economics After the Crisis: Objectives and Means*. MIT Press.
Turner, P. A., Field, C. B., Lobell, D. B., Sanchez, D. L., & Mach, K. J. (2018). Unprecedented rates of land-use transformation in modelled climate change mitigation pathways. *Nature Sustainability*, *1*(5), 240–245. https://doi.org/10.1038/s41893-018-0063-7
UNEP. (2011). *Towards a Green Economy: Pathways to Sustainable Development and Poverty Eradication*. https://doi.org/10.1063/1.3159605
UNFCCC. (2019). *The Paris Agreement*. https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement
Usher, W. (2016). *The Value of Learning about Critical Energy System Uncertainties* (Issue September). http://discovery.ucl.ac.uk/1504608/%5Cnhttp://discovery.ucl.ac.uk/id/eprint/1504608
Van der Sluijs, J. P., Potting, J., Risbey, J., van Vuuren, D., de Vries, B., Beusen, A., Heuberger, P., Corral Quintana, S., Funtowicz, S. O., & Kloprogge, P. (2002). Uncertainty assessment of the IMAGE/TIMER B1 CO2 emissions scenario, using the NUSAP method. In *Report commissioned by the Netherlands National Research Program on global Air Pollution and Climate Change, RIVM, Bilthoven, The Netherlands* (Issue 954267). http://www.nusap.net
van Vliet, J., den Elzen, M. G. J., & van Vuuren, D. P. (2009). Meeting radiative forcing targets under delayed participation. *Energy Economics*, *31*(SUPPL. 2), S152–S162. https://doi.org/10.1016/j.eneco.2009.06.010
Veblen, T. (1898). Why is economics not an evolutionary science? *The Quarterly Journal of Economics*, *12*(4), 373–397. https://doi.org/10.1177/0309132512463300
Vries, B. J. M. De, Vuuren, D. P. Van, Elzen, M. G. J. Den, & Janssen, M. A. (2001). *The targets image energy regional (timer) model*.
Waisman, H., Guivarch, C., Grazi, F., & Hourcade, J. C. (2012). The Imaclim-R model: Infrastructures, technical inertia and the costs of low carbon futures under imperfect foresight. *Climatic Change*, *114*(1), 101–120. https://doi.org/10.1007/s10584-011-0387-z
Weinberg, a M. (1972). Science and trans-science. *Science (New York, N.Y.)*, *177*(4045), 211. https://doi.org/10.1126/science.177.4045.211
Weitzman, M. L. (2009). On Modeling and Interpreting the Economics of Catastrophic Climate Change. *Review of Economics and Statistics*, *91*(1), 1–19. https://doi.org/10.1162/rest.91.1.1
Wilson, C., Grubler, A., Bauer, N., Krey, V., & Riahi, K. (2013). Future capacity growth of energy technologies: Are scenarios consistent with historical evidence? *Climatic Change*, *118*(2), 381–395. https://doi.org/10.1007/s10584-012-0618-y
Wing, I. S., Calvin, K., Danzer, K., & Fisher-Vanden, K. (2011). *Phoenix Model Documentation* (Issue August). http://www.globalchange.umd.edu/data/models/phx_documentation_august_2011.pdf
Wolf, S., Schutze, F., & Jaeger, C. C. (2016). *The Possibility of Green Growth in Climate Policy Analysis Models – a Survey*.
Wolf, S., Schütze, F., & Jaeger, C. C. (2016). Balance or synergies between environment and Economy-A note on model structures. *Sustainability (Switzerland)*, *8*(8), 1–11. https://doi.org/10.3390/su8080761
World Bank. (2012). *Inclusive Green Growth: The Pathway to Sustainable Development*. https://doi.org/10.1596/978-0-8213-9551-6
Wright, J. (2018). *Pluralism and social epistemology in economics* [Cambridge].

https://www.repository.cam.ac.uk/bitstream/handle/1810/290423/J Wright - Pluralism and social epistemology in economics.pdf?sequence=1&isAllowed=y
Yamamoto, H., Sugiyama, M., & Tsutsui, J. (2014). Role of end-use technologies in long-term GHG reduction scenarios developed with the BET model. *Climatic Change*, *123*(3–4), 583–596. https://doi.org/10.1007/s10584-013-0938-6
Yue, X., Pye, S., DeCarolis, J., Li, F. G. N., Rogan, F., & Gallachóir, B. (2018). A review of approaches to uncertainty assessment in energy system optimization models. *Energy Strategy Reviews*, *21*(July 2017), 204–217. https://doi.org/10.1016/j.esr.2018.06.003
Zhang, Z., & Folmer, H. (1998). Economic modelling approaches to cost estimates for the control of carbon dioxide emissions. *Energy Economics*, *20*(1), 101–120. https://doi.org/10.1016/S0140-9883(97)00019-4

A Links between IAMs

A.1 MERGE group (BET, REMIND, WITCH, MERGE, MERGE-ETL)

The **BET** (Basic Energy systems, Economy, Environment, and End-use Technology) model is the most recent addition to the 14 IAMs. In introducing the model for the first time, Yamamoto et al. (2014, p. 584), tell us that BET "is strongly influenced by MERGE (Manne et al. 1995; Richels and Blanford 2008) and MARKAL-MACRO (Loulou et al. 2004), which is closely related with TIAM". In fact, they continue, "roughly speaking, the BET model can be summed up as "a MERGE with advanced, electric end-use technologies" or "a global MARKAL-MACRO with limited technologies"" (2014, p. 585). There are indications that MERGE did not only just serve as a theoretical foundation, but also that the computer code itself was utilised in the design of BET. As the acknowledgments state, "we greatly appreciate the kindness of the MERGE group to make a version of the code available online, which helped us develop the BET model" (2014, p. 595).

REMIND (Regionalized Model of Investments and Development) is the most influential as well as one of the newest IAMs. The first version (called REMIND-R) was only introduced in 2010 (Leimbach et al., 2010). Leimbach et al. (2010) tell us that "with MERGE and WITCH, REMIND-R shares the same intertemporal structure, but is distinguished from both by a higher degree of technological resolution in the energy sector" (Leimbach et al., 2010, p. 157). Luderer et al. (2013, p. 2) similarly tells us that "in terms of its macro-economic formulation, REMIND resembles well-known energy-economy-climate models such as RICE (Nordhaus and Yang 1996) and MERGE (Manne et al. 1995). However, REMIND features a higher level of detail in the representation of energy-system technologies, trade, and global capital markets".

The **WITCH** (World Induced Technical Hybrid) model is also a recent and highly influential IAM, presented in a highly cited paper by Bosetti et al. in 2006. Bosetti et al. (2006) compares WITCH to three other IAMs: MERGE, RICE, and MIND (a forerunner to REMIND). More specifically, WITCH is compared to MERGE in the following way "MERGE (Manne, Mendelsohn and Richels, 1995) links a simple top-down model to a bottom-up part that returns the cost of energy; in contrast, WITCH is a single model that represents the energy sector within the economy, and therefore chooses the energy technology investment paths coherently with the optimal growth structure" (Bosetti, Carraro, Galeotti, et al., 2006, p. 16). With respect to MIND, "WITCH possesses richer technological detail, differentiates the electric and non-electric energy uses and is a regional mode" (Bosetti, Carraro, Galeotti, et al., 2006, p. 16). (When it comes to RICE "WITCH shares a game set-up similar to that in RICE (Nordhaus and Boyer, 2000), but departs from the stylized representation of the energy sector by featuring richer technological detail, technical change, natural resource depletion etc" (Bosetti, Carraro, Galeotti, et al., 2006, p. 16)).

The relationship between MERGE and WITCH is the least clear one. Implicitly, however, by stating only that what distinguishes WITCH from MERGE is the way in which the top-down and bottom-up components are combined, similarities in other aspects of the two models can be inferred.

The starting point for this group is **MERGE** (Model for Evaluating Regional and Global Effects of GHG reduction policies), which was introduced by Manne, Mendelsohn and Richels in 1995 (Manne et al., 1995). MERGE was constructed by combining an existing model used to estimate the cost of emissions constraints, Global 2200, with a climate and a damage assessment module[94]. Global 2200, again, was presented (as Global 2100) in the highly influential book, *Buying Greenhouse Insurance*, by Manne and Richels in 1992 (Manne & Richels, 1992). Global 2100, again, was designed by linking to existing models, MACRO and ETA. It presented a "two-way linkage between a top-down model of economic growth and energy demands (MACRO) and a bottom-up model for energy technology assessment (ETA)" (Stanford University, 2019). Global 2100/2200 is therefore sometimes referred to as ETA-MACRO.

MERGE-ETL is simply a modified version of MERGE, introduced in 2003, which includes endogenous technical learning (hence the addition of the term, ETL) (Kypreos & Bahn, 2003)[95].

In summary, the structure of BET, REMIND, and MERGE-ETL resemble the structure of MERGE, but with increasing technological detail and/or an endogenous representation of technical change.

A.2 MESSAGE-MACRO

The label **MESSAGE** is used both to denote the MESSAGE energy system models and the entire IAM framework developed at the International Institute for Applied Systems Analysis (IIASA), in which MESSAGE form a crucial component. MESSAGE itself has part of this framework from the very beginning, with the first version dating back to 1979 (Agnew et al., 1979b). This makes MESSAGE the oldest models in the AR5 ensemble. The version of MESSAGE used in AR5, however, should really be called MESSAGE-MACRO. This IAM was constructed by linking the energy system model,

[94] Allowing for cost-benefit analysis of climate change.

[95] Essentially, the difference between MERGE and MERGE-ETL is the introduction of R&D as a decision variable in the optimisation.

MESSAGE, with the macroeconomic model, MACRO, in a highly cited journal paper from 2000 (Messner & Schrattenholzer, 2000).

The **MACRO** model (of MESSAGE-MACRO) is the same model as the MACRO model in ETA-MACRO (Manne et al., 1995). The link between MERGE and MESSAGE, however, started before MESSAGE-MACRO. We find that "prior to the development of MESSAGE–MACRO and its inclusion in IIASA's Integrated Assessment Scheme, two separate models, MESSAGE and 11R, played the role that is now fulfilled by the linked modules" (Messner & Schrattenholzer, 2000, p. 269). 11R, again, is "a model building on the Global 2100 model by Manne and Richels" (Messner & Schrattenholzer, 2000, p. 269). Overall, "the form of MACRO used in the IIASA IAM framework is derived from a long series of models by Manne and Richels" (Fricko et al., 2017, p. 258).

According to Messner & Schrattenholzer (2000) MESSAGE–MACRO (which will be discussed below) also has much in common with MARKAL-MACRO. More specifically, "the main difference between the two approaches is that MARKAL–MACRO is a fully integrated single model, whereas MESSAGE–MACRO is solved by running each part separately and iterating their inputs until consistency between the macroeconomic part and the energy part is reached" (p. 270).

A.3 TIAM-World

TIAM stands for the 'TIMES Integrated Assessment Model' and *TIAM-World* is the global multiregional incarnation of this (Labriet et al., 2012; Richard Loulou & Labriet, 2008). TIMES, again, which is short for 'The Integrated MARKAL-EFOM System', "was developed as a successor of the MARKAL (Fishbone and Abilock 1981; Fishbone et al. 1983; Berger et al. 1992), and EFOM (Finon 1974; van der Voort et al. 1984) bottom-up energy models" (Loulou and Labriet, 2008)[96]. EFOM dates back to 1974 (Finon, 1974), but no longer appears to be around. MARKAL, on the other hand, was developed over a period of almost two decades by the Energy Technology Systems Analysis Programme (ETSAP) of the International Energy Agency, starting in the late 1970s (Taylor et al., 2014). Loulou et al. (2004, p. 389) tell us that "a precursor of [MARKAL-MACRO] is ETA-MACRO (Manne and Richels 1992), where the ETA module was however a less detailed energy supply module than MARKAL is." MARKAL is today seen as one of the most successful energy models of recent decades, with the number of users currently at 77 institutions in 37 countries (ETSAP, 2019).

---

[96] Thus, if one were to spell out the full name of TIAM it would be 'The Integrated MARKAL-EFOM System Integrated Assessment Model'.

A.4 AIM-Enduse – GCAM link

The **AIM-Enduse** model is an extension of the Asian-Pacific Integrated Model (AIM), which was developed at the National Institute for Environmental Studies (NIES) in Japan between 1991 and 1994. Nakicenovic et al. (2000, p. 336) tell us that parts of AIM is based on ERB. The first source describing AIM-Enduse is a working paper from 1995 (Kainuma et al., 1995). Although no explicit links to MESSAGE are stated (and therefore no links are drawn in the model family tree), it is worth noting that this working paper was published at IIASA, which is where the MESSAGE model was and still is developed.

GCAM (Global Change Assessment Model) is one of the most influential and oldest IAMs, developed at The Joint Global Change Research Institute (JGCRI). GCAM was first known as the Edmonds-Reilly (and subsequently the Edmonds-Reilly-Barnes (ERB)) model, which was presented in 1983 (Edmonds & Reilly, 1983b, 1983c, 1983a). The model was renamed MiniCAM in the mid-1990s and GCAM in the mid-2000s (GCAM v5.1 Documentation, 2019).

Wing et al. (2011) inform us that "**Phoenix** is a re-design of the Second Generation Model (SGM) produced by The Joint Global Change Research Institute, which was developed as a *complement* to the "first generation model" known as MiniCAM" (which later became GCAM). SGM "is a process-level version of the ERB" (Brenkert et al., 2003, p. 13). In other words, GCAM and Phoenix share the same roots in the ERB model (which is not surprising, seeing that the two models were developed at the same institution).

A.5 POLES – IMAGE link

The **POLES** (Prospective Outlook on Long-term Energy Systems) model was initially developed in the 1990s in France before being transferred to the EU Joint Research Centre (JRC) (Despres, 2018). The model was first presented in 1990 (Lesourd et al., 1996). The limited documentation does not state any links to other AR5 models (or forerunners of). According to Lesourd (1996), "the structure of the model and its operating logic are based on two key concepts established by H. A. Simon". The fact that Lesourd mentions the intellectual roots, but no other IAMs, however, indicates that POLES was developed relatively independently of other IAMs.

Much like MESSAGE, **IMAGE** (Integrated Model to Assess the Global Environment) also represents a modelling framework consisting of several sub-models. The history of IMAGE goes back to 1985, when Rotmans started the work on the initial prototype, leading to the first publication in 1990 (Rotmans, 1990). According to the IMAGE 2.4 model manual, "IMAGE 1.0 was among the first pioneering examples of Integrated Assessment Models addressing climate change"(Bouwman et al., 2006, p. 17). Rotmans (1990) mentions two similar models, but none of these are included in AR5 (or referred to by other models in AR5)[97]. This also signals independence from other early IAMs.

The version of IMAGE used in AR5 in reality consists of three models: IMAGE, TIMER, and FAIR (Bouwman et al., 2006). Of these, TIMER (The IMage Energy Regional model) covers the energy system and is based on the TIME model, which was first presented by Vries et al. (1995). According to Vries et al. (2001, p. 11) "a model which is in various aspects similar to the TIMER-model is the POLES-model". The FAIR (Framework to Assess International Regimes) model is a global climate policy model that was first introduced by Den Elzen and Lucas in 2003.

In summary, IMAGE and POLES appears to have been developed independently around the same time. The only connection between the two is the energy system model, TIMER, which appears to have been inspired by POLES and incorporated into the IMAGE framework later on.

A.6 IMACLIM

IMACLIM, developed at the Center for International Research on Environment and Development (CIRED) in France, was first introduced in 1991 (Baron & Salles, 1991). IMACLIM was "an adaptation of the Chandler (1990) model to France" (Beaumais & Zagame, 1993, p. 118). While Beaumais and Zagame argue that IMACLIM differed from all other models developed in France at the time, which were "econometric neo-Keynesian models" (Beaumais & Zagame, 1993, p. 113), they do not mention any of the other models included in AR5. Hourcade (1993) also argues that IMACLIM tackles issues that existing models are not able to. Overall, however, fairly little is written (in English) on the history of IMACLIM and on how it relates to other models. Based on what we have, however, IMACLIM appears to have been developed relatively independently of others IAMs.

A.7 DNE+21

The **DNE21+** model, from The Research Institute of Innovative Technology for the Earth (RITE) in Japan, was introduced by Sano et al. (2005). DNE is short for "Dynamic New Earth". The forerunners to DNE21+ are LDNE21 and DNE21 (described by Yamaji et al. (2000) and Fujii & Yamaji (1998) respectively). According to Fuji & Yamaji (1998), they "built a new global energy system model, Dynamic New Earth 21, based on the New Earth 21 model developed previously" (Fujii & Yamaji,

---

[97] The Model of Warming Commitment (MWC) of the World Resource Institute (Mintzer, 1987), and the Atmospheric Stabilization Framework (ASF) of the EPA (EPA, 1989).

1998, p. 114). The New Earth 21 model dates back to 1993. No links are found to other models, but we note that the available documentation for this model (and its forerunners) is fairly limited.

B IAM Frameworks

This appendix describes how the model documentation for the 14 IAMs present the underlying model frameworks.

B.1 Ramsey-type optimal growth models

In Global 2100, which formed the basis of **MERGE**, "a Ramsey model is employed for the determination of savings and investment through a discounted utility maximand" Stanford (2019). **WITCH** is described as "a Ramsey-type neoclassical optimal growth hybrid model" (Bosetti, Carraro, Galeotti, et al., 2006, p. 15) and **BET** is described as "a multi-regional, global model based on Ramsey's optimal growth theory" (Yamomoto et al., 2014). Similarly, we find that the "macro-economic core of **REMIND** is a Ramsey-type optimal growth model in which intertemporal global welfare is optimized subject to equilibrium constraints" (web site). (**MERGE-ETL** is based on MERGE and shares the same basic framework.)

The reason why **MESSAGE** is described as a *general* equilibrium model in AR5 is because it is coupled to MACRO (to form MESSAGE-MACRO). Given that MACRO stems directly from MERGE, MACRO is also based on the same framework as MERGE (and Global 2100). We can thus consider MESSAGE-MACRO part of the MERGE family. In presenting MESSAGE-MACRO, Messner & Schrattenholzer (2000, p. 270) described MACRO as "a macroeconomic model maximizing the intertemporal utility function of a single representative producer-consumer in each world region. The optimization result is a sequence of optimal savings, investment, and consumption decisions."

MESSAGE itself (without MACRO) is the energy system model in the IIASA IAM framework. MESSAGE alone is described as "a dynamic linear programme for comparing alternative existing and new energy supply technologies" (Agnew et al., 1979b, p. 4). Messner & Schrattenholzer (2000) use the slightly different term, "dynamic systems engineering optimization model", to describe the same framework in 2000 (p. 270).

B.2 Energy system optimisation models (ESOMs)

TIAM-World is based on TIMES, which again is based on MARKAL. All of these are described as "technology explicit, dynamic partial equilibrium models of energy markets" (R Loulou et al., 2016, p. 135). In TIMES (and MARKAL) "the equilibrium is obtained by maximizing the total surplus of consumers and suppliers via Linear Programming, while minimizing total discounted energy system cost" (R Loulou et al., 2016, p. 135). By maximising total surplus, TIAM-World "computes a dynamic inter-temporal partial equilibrium on worldwide *energy and emission markets*" (Website, my italics). **DNE21+** is also described as a linear programming model, in which net energy system costs are minimised (RITE web site).

B.3 CGE models

IMACLIM[98] is described as a "recursive dynamic, multi-region and multi-sector hybrid CGE model" (Bibas and Méjean, 2014, p. 734). **Phoenix** is described as "a recursive dynamic computable general equilibrium (CGE) model" (Wing et al., 2011).

B.4 Energy simulation models

GCAM is described as a "dynamic-recursive market equilibrium model" (Kriegler et al., 2015a, p. 7). Moreover, "the core operating principle for GCAM is that of market equilibrium" (Model Overview Wiki). The reason why it's considered a partial rather than general equilibrium model is because "only markets for certain goods such as energy and agricultural goods are represented and cleared" (Lurz et al., 2006, pp. 72–73). **AIM-Enduse** represents one of several emissions sub-modules in the AIM model. It is a recursive dynamic (Hanaoka et al., 2015) bottom-up energy model that "focuses on the end-use technology selection in energy consumption as well as energy production" (Kainuma et al., 2003, p. 8). According to the web site, "**POLES** is a world energy-economy partial equilibrium simulation model of the energy sector" (Website 2), which follows a recursive dynamic.

IMAGE self-declares as closer to earth systems models than any other IAMs. The focus of IMAGE is on the representation of physical processes and geographical detail, rather than economic processes and feedback (Stehfest et al., 2014). For the purpose of investigating climate change mitigation strategies in AR5 (and AR4) IMAGE is linked to FAIR and TIMER to create IMAGE-TIMER-FAIR. In order to understand why IMAGE is categorised as a recursive-dynamic partial equilibrium model, we need to look to the energy system module, TIMER. According to Lucas et al. (2013), "TIMER is a recursive dynamic global energy-system model that describes the long-term dynamics of the production and consumption of energy" (p. 1033). According to the web site "the focus is on dynamic relationships in the energy system, such as inertia and learning-by-doing in capital stocks, depletion of the resource base and trade between regions". FAIR is included "for inter-temporal optimisation of mitigation

[98] IMACLIM is available in a static version (IMACLIM-S) and a recursive version (IMACLIM-R). It is the latter version that is used in AR5.

strategies”. In this set-up, long-term reduction strategies are determined by minimising the cumulative discounted mitigation costs (the latter being calculated by TIMER) (Stehfest et al., 2014).

C AR5 general equilibrium IAM publications

This appendix lists the AR5 scenario publications that were used in section 4.4 to identify potential reasons why the 20 general equilibrium IAMs in AR5 do not generate a single net negative cost result. 16 of the publications are model inter-comparison overview sources presenting results from multiple IAMs. As such, this appendix also shows what general equilibrium IAMs were part of what model inter-comparison exercise. Two general equilibrium IAMs in AR5 were not part of any of the model inter-comparison exercises. For these two IAMs, the specific publications that presented the AR5 scenarios were used instead.

Table C.1 AR5 scenario publications covering scenarios generated by general equilibrium IAMs.

	General equilibrium IAMs used in publications
Model inter-comparison overview publications	
AMPERE (Kriegler et al., 2015b; Riahi et al., 2015b)	GEM-E3-ICCS, MERGE-ETL_2011, MESSAGE V.4, REMIND 1.5, WITCH_AMPERE, WorldScan2
EMF 27 (G. J. Blanford et al., 2014; Krey et al., 2014; Kriegler et al., 2014)	BET 1.5, EC-IAM 2012, FARM_3.0, GRAPE_ver1998, IMACLIM v1.1, MERGE_EMF27, MESSAGE V.4, Phoenix 2012.4, REMIND 1.5
RoSE (Bauer et al., 2016; Calvin et al., 2016; Chen et al., 2016; De Cian et al., 2016; Luderer et al., 2016)	REMIND 1.4, WITCH_ROSE
LIMITS (Kriegler et al., 2013; Tavoni et al., 2014)	MESSAGE V.4, REMIND 1.5, WITCH_LIMITS
AME (Calvin et al., 2012)	MESSAGE V.3, REMIND 1.3, GRAPE_ver1998, MARIA23_org, MERGE_AME, Phoenix 2012.4, WITCH_AME, GTEMREF32, iPETS_1.2.0
EMF 22 (Clarke et al., 2009)	MERGE_EMF22, MESSAGE V.1, SGM_EMF22, WITCH_EMF22
RECIPE (Luderer et al., 2012)	IMACLIM, REMIND, WITCH

ADAM (Edenhofer, Knopf, Leimbach, & Bauer, 2010)	MERGE-ETL
Individual AR5 scenario publications	
MESSAGE RCP 8.5 scenarios (Riahi et al., 2011)	MESSAGE V.2
IGSM scenarios (Prinn et al., 2011)	IGSM

D Comparison of parameter ranges used in Chapter 6 with values in the literature

D.1 Learning rates

Table D.1 Learning rate ranges in FTT:Power and in the literature

Technology	**FTT:Power ranges**	**Samadi** (2018)	**Rubin et al.** (2015)
Nuclear	5.8 (2.9, 8.7)	(-25, 10)	(-38*, 6)
Oil	1.0 (0.5, 1.5)		
Coal	3.0 (1.5, 4.5)	(-5, 5)	(5.6, 12)
Coal + CCS	5.0 (2.5, 7.5)		(1.1, 9.9)
IGCC	3.0 (1.5, 4.5)		(2.5, 16)
IGCC + CCS	5.0 (2.5, 7.5)		(2.5, 20)
CCGT	4.0 (2.0, 6.0)	(2, 15)	(-11, 34)
CCGT + CCS	5.0 (2.5, 7.5)		(2, 7)
Solid Biomass	5.0 (2.5, 7.5)		(0, 24)
S Biomass CCS	7.0 (3.5, 10.5)		
BIGCC	5.0 (2.5, 7.5)		
BIGCC + CCS	7.0 (3.5, 10.5)		
Biogas	5.0 (2.5, 7.5)		
Biogas + CCS	7.0 (3.5, 10.5)		
Tidal	1.4 (0.7, 2.1)		
Large Hydro	1.4 (0.7, 2.1)		(1.4, 1.4)
Onshore	7.0 (3.5, 10.5)	(-3, 12)	(-11, 32)
Offshore	9.0 (4.5, 13.5)	(-5, 10)	(5, 19)
Solar PV	17.0 (8.5, 25.5)	(8, 23)	(10, 47)
CSP	10.0 (5.0, 15.0)	(3, 12)	
Geothermal	5.0 (2.5, 7.5)		
Wave	14.0 (7.0, 21.0)		
Fuel Cells	15.0 (7.5, 22.5)		
CHP	3.0 (1.5, 4.5)		

* Rubin (2015) notes these values should not be translated into learning rates due to the many factors that may lead nuclear power to become more expensive rather than less.

The high ends of the learning rates, which result from multiplying the FTT:Power default learning rates by 150%, are close to those reported by Samadi (2018) for nuclear, coal, onshore wind, offshore wind, solar PV, and CSP. Only for offshore wind and for solar PV are the high ends of the learning rates slightly higher than what is reported by Samadi (2018). Rubin et al. (2015) however, report higher learning rates for both of these technologies.

The low ends of the learning rates, which result from multiplying the FTT:Power default learning rates by 50%, are close to those reported by Samadi (2018) for gas (CCGT), solar PV, and CSP, and conservative for nuclear, coal, and onshore and offshore wind. In particular, the negative learning rates reported by Samadi (2018) are not captured by varying the FTT:Power default learning rates by ±50%.

Overall, the learning rates that arise from varying the FTT:Power default values by ±50% are either similar to or conservative compared to those reported by Samadi (2018) and Rubin et al. (2015).

D.2 Lifetimes

Krey et al. (2019b) provide a review of techno-economic assumptions in the electricity sector in fifteen IAMs. Among other things, they review lifetime ranges.

Table D.2 Lifetime ranges in FTT:Power and in the literature

Technology	FTT:Power ranges	Krey et al.
Nuclear	60 (30, 90)	(40,60)
Oil	40 (20, 60)	
Coal	40 (20, 60)	(30,60)
Coal + CCS	40 (20, 60)	(30,60)
IGCC	40 (20, 60)	
IGCC + CCS	40 (20, 60)	
CCGT	30 (15, 45)	(25,45)
CCGT + CCS	30 (15, 45)	(25,45)
Solid Biomass	40 (20, 60)	(20,60)
S Biomass CCS	40 (20, 60)	(20,60)
BIGCC	40 (20, 60)	
BIGCC + CCS	40 (20, 60)	
Biogas	30 (15, 45)	
Biogas + CCS	30 (15, 45)	
Tidal	80 (40, 120)	
Large Hydro	80 (40, 120)	(40, inf)
Onshore	25 (12.5, 37.5)	(20,31)
Offshore	25 (12.5, 37.5)	(20,31)
Solar PV	25 (12.5, 37.5)	(20,30)

CSP	25 (12.5, 37.5)	(20,30)
Geothermal	40 (20, 60)	(30,40)
Wave	20 (10, 30)	
Fuel Cells	20 (10, 30)	
CHP	40 (20, 60)	

Inf means that the lifetime of a technology for modelling purposes is unlimited

Several of the lifetime ranges found by Krey et al. (2019b) are remarkably close to the ranges obtained by varying FTT:Power default values by ±50%. The lower ends are, however, slightly lower in FTT:Power compared to what is reported in Krey et al. (2019b), and for nuclear, wind, solar PV, geothermal, and CSP the high ends in FTT:Power are also somewhat higher (the role of geothermal and CSP in the scenarios, however, are not significant). Note, however, that with lifetimes – just as with build times – it is the relative difference between technologies that matter (varying the absolute values is equivalent to varying the overall scaling factor). Overall, the ranges reported in Krey et al. (2019b) are somewhat narrower, but still close to the ranges obtained by varying FTT:Power default values by ±50%. Given that Krey et al. (2019b) does not provide an exhaustive review of lifetime ranges in the literature, the FTT:Power ranges do not appear unreasonable.

D.3 Investor discount rates

The default investor discount rate (hurdle rate) in FTT:Power is 10% for all technologies. In a sensitivity analysis of the widely used MARKAL model, Johnson et al. (2006) varied the hurdle rate for new electricity generation technologies between 5% and 20%. This range was based on modeler judgment. Labriet et al. (2012) also use hurdle rates between 5% and 20% in their exploration of the impacts of technology and climate uncertainties on optimal pathways generated by TIAM-World. The IEA's "Projected cost of generating electricity" (IEA & NEA, 2015) uses hurdle rates of 5% and 10% to calculate the costs of power technologies. Based on this, the range (0.05, 0.20) obtained by varying FTT:Power default values by ±50% appears reasonable.

E Remaining FTT:Power parameters

This appendix lists FTT:Power parameters that are not included in the sensitivity analysis.

Table E.1 Parameters used to compute LCOEs, shares, capacities, and emissions in FTT:Power

		Investment	std	Fuel	std	O&M	std	Load Factor	Type	Efficiency	Resource	Emissions
		$/kW	$/MWh	$/MWh	$/MWh	$/MWh	$/MWh		0;1;2;3	%	Efficiency	tCO2/GWh
1	Nuclear	4896	1525.05	9.6	2.331842	11	6.150083	0.95	1	1	1	0
2	Oil	1227.845	1033.628	223.6639	239.5182	22.126	5.694103	0.85	1	0.45	0.272617	751.8497
3	Coal	2292.949	775.0097	25.61937	11.22748	7.413898	6.022754	0.85	1	0.42	0.41682	998.3275
4	Coal + CCS	4224.692	1172.546	22.43323	10.22969	15.01657	4.554923	0.85	1	0.37	0.367199	99.83275
5	IGCC	3829.065	1705.944	20.05085	1.570261	10.0901	1.50928	0.85	1	0.42	0.41682	998.3275
6	IGCC + CCS	4521.142	1523.045	19.9646	7.502858	12.87135	0.521059	0.85	1	0.37	0.367199	99.83275
7	CCGT	1067	336.7544	66.45885	16.52231	5.821213	2.797362	0.85	1	0.57	0.513211	504.2741
8	CCGT + CCS	2446.527	520.6255	71.19667	1.471167	6.419504	0.403999	0.85	1	0.47	0.423174	50.42741
9	Solid Biomass	4007	2587.467	93.24	72.93987	18.55	26.53242	0.85	2	0.42	0.310116	0
10	S Biomass CCS	5938.743	2985.004	93.24	72.93987	18.55	26.53242	0.85	2	0.37	0.273197	-980.757
11	BIGCC	3829.065	1705.944	93.24	72.93987	10.0901	1.50928	0.85	2	0.42	0.310116	0
12	BIGCC + CCS	4521.142	1523.045	93.24	72.93987	12.87135	0.521059	0.85	2	0.37	0.273197	-980.757
13	Biogas	3733	3519.629	0	36.61555	60.52	5.839101	0.85	3	0.57	0.513211	0
14	Biogas + CCS	5112.527	3703.5	0	36.61555	60.52	5.839101	0.85	3	0.47	0.423174	-376.391
15	Tidal	2782.5	3538.984	0	0	38.4	6.451792	0.3	3	1	1	0
16	Large Hydro	2492.5	2499.96	0	0	9.855	10.42698	0.4	3	1	1	0
17	Onshore	1841	443.4874	0	0	21.38	8.673325	0.265	0	1	1	0
18	Offshore	5000	579.5776	0	0	40.71	19.82348	0.39	0	1	1	0
19	Solar PV	1833.5	552.8973	0	0	22.795	15.56944	0.16	0	1	1	0
20	CSP	4901	1859.097	0	0	17.38	22.09532	0.32	0	0.2	0.2	0
21	Geothermal	5822.5	2036.632	0	0	17.275	34.09743	0.85	3	1	1	0
22	Wave	5142.072	2414.849	0	0	55.9106	36.58099	0.455	0	1	1	0
23	Fuel Cells	5884.815	5459	58.70801	54.56	53.6953	49.81	0.85	1	0.8	0.720296	359.2953
24	CHP	2000	4358.279	65.74	15.20814	15.93	31.84846	0.85	1	0.8	0.720296	359.2953

Table E.2 Spillover learning matrix (see Mercure (2011, 2012) for the full set of equations describing the computation of learning effects in FTT:Power).

	1 means technologies are the same; 0 they are unrelated	Nuc	Oil	Coal	Coal + CCS	IGCC	IGCC + CCS	CCGT	CCGT + CCS	S Bio	S Bio CCS	BIGCC	BIGCC + CCS	Biogas	Biogas + CCS	S Hydro	L Hydro	Onshore	Offshore	Solar PV	CSP	Geotherm	Wave	Fuel Cells	CHP
1	Nuclear	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Oil	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Coal	0	0	1	0.75	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Coal + CCS	0	0	0.75	1	0	1	0	1	0	0.75	0	1	0	1	0	0	0	0	0	0	0	0	0	0
5	IGCC	0	0	0	0	1	0.75	0.5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	IGCC + CCS	0	0	0	1	0.75	1	0	0.75	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
7	CCGT	0	0	0	0	0.5	0	1	0.75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	CCGT + CCS	0	0	0	1	0	0.75	0.75	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
9	Solid Biomass	0	0	0.5	0	0	0	0	0	1	0.75	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	S Biomass CCS	0	0	0	0.75	0	1	0	1	0.75	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
11	BIGCC	0	0	0	0	1	0	0	0	0	0	1	0.75	0	0	0	0	0	0	0	0	0	0	0	0
12	BIGCC + CCS	0	0	0	1	0	1	0	1	0	1	0.75	1	0	1	0	0	0	0	0	0	0	0	0	0
13	Biogas	0	0	0	0	0	0	0	0	0	0	0	0	1	0.75	0	0	0	0	0	0	0	0	0	0
14	Biogas + CCS	0	0	0	1	0	1	0	1	0	1	0	1	0.75	1	0	0	0	0	0	0	0	0	0	0
15	Tidal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
16	Large Hydro	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
17	Onshore	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0
18	Offshore	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	1	0	0	0	0	0	0
19	Solar PV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.1	0	0	0	0
20	CSP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	1	0	0	0	0
21	Geothermal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
22	Wave	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
23	Fuel Cells	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
24	CHP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table E.3 Capacity factors for each technology and region in FTT:Power.

		Belgium	Denmark	Germany	Greece	Spain	France	Ireland	Italy	Luxembourg	Netherlands	Austria	Portugal	Finland	Sweden	UK	Czech Republic	Estonia	Cyprus	Latvia	Lithuania
1	Nuclear	0.89	0.85	0.83	0.85	0.91	0.79	0.85	0.85	0.85	0.93	0.85	0.85	0.98	0.82	0.85	0.81	0.85	0.85	0.85	0.95
2	Oil	0.30	0.30	0.38	0.30	0.34	0.30	0.30	0.30	0.30	0.30	0.49	0.43	0.30	0.30	0.30	0.26	0.21	0.60	0.30	0.30
3	Coal	0.60	0.60	0.48	0.63	0.60	0.60	0.60	0.60	0.85	0.60	0.60	0.32	0.60	0.60	0.44	0.55	0.44	0.85	0.60	0.60
4	Coal + CCS	0.60	0.60	0.48	0.63	0.60	0.60	0.60	0.60	0.85	0.60	0.60	0.32	0.60	0.60	0.44	0.55	0.44	0.85	0.60	0.60
5	IGCC	0.60	0.60	0.48	0.63	0.60	0.60	0.60	0.60	0.85	0.60	0.60	0.32	0.60	0.60	0.44	0.55	0.44	0.85	0.60	0.60
6	IGCC + CCS	0.60	0.60	0.48	0.63	0.60	0.60	0.60	0.60	0.85	0.60	0.60	0.32	0.60	0.60	0.44	0.55	0.44	0.85	0.60	0.60
7	CCGT	0.55	0.48	0.72	0.50	0.55	0.64	0.62	0.48	0.68	0.80	0.37	0.66	0.37	0.55	0.66	0.55	0.55	0.85	0.55	0.55
8	CCGT + CCS	0.55	0.48	0.72	0.50	0.55	0.64	0.62	0.48	0.68	0.80	0.37	0.66	0.37	0.55	0.66	0.55	0.55	0.85	0.55	0.55
9	Solid Biomass	0.57	0.45	0.47	0.50	0.55	0.48	0.50	0.47	0.47	0.52	0.50	0.61	0.68	0.39	0.37	0.47	0.50	0.85	0.50	0.29
10	S Biomass CCS	0.57	0.45	0.47	0.50	0.55	0.48	0.50	0.47	0.47	0.52	0.50	0.61	0.68	0.39	0.37	0.47	0.50	0.85	0.50	0.29
11	BIGCC	0.57	0.45	0.47	0.50	0.55	0.48	0.50	0.47	0.47	0.52	0.50	0.61	0.68	0.39	0.37	0.47	0.50	0.85	0.50	0.29
12	BIGCC + CCS	0.57	0.45	0.47	0.50	0.55	0.48	0.50	0.47	0.47	0.52	0.50	0.61	0.68	0.39	0.37	0.47	0.50	0.85	0.50	0.29
13	Biogas	0.55	0.48	0.72	0.50	0.50	0.64	0.62	0.48	0.68	0.80	0.37	0.66	0.37	0.50	0.66	0.50	0.50	0.85	0.50	0.50
14	Biogas + CCS	0.55	0.48	0.72	0.50	0.50	0.64	0.62	0.48	0.68	0.80	0.37	0.66	0.37	0.50	0.66	0.50	0.50	0.85	0.50	0.50
15	Tidal	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
16	Large Hydro	0.42	0.33	0.68	0.40	0.20	0.40	0.46	0.35	0.44	0.31	0.41	0.19	0.63	0.48	0.36	0.22	0.64	0.85	0.23	0.40
17	Onshore	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
18	Offshore	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
19	Solar PV	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
20	CSP	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
21	Geothermal	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
22	Wave	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
23	Fuel Cells	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
24	CHP	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

Table E.3 cont.

		Hungary	Malta	Poland	Slovenia	Slovakia	Bulgaria	Romania	Norway	Switzerland	Iceland	Croatia	Turkey	Macedonia	USA	Japan	Canada	Australia	New Zealand	Russian Federation
1	Nuclear	0.87	0.85	0.85	0.85	0.87	0.95	0.91	0.85	0.98	0.85	0.85	0.85	0.85	0.95	0.85	0.85	0.85	0.85	0.82
2	Oil	0.30	0.30	0.30	0.30	0.30	0.69	0.42	0.13	0.30	0.20	0.30	0.66	0.85	0.11	0.32	0.25	0.19	0.20	0.11
3	Coal	0.60	0.85	0.55	0.70	0.60	0.58	0.60	0.60	0.85	0.85	0.60	0.62	0.60	0.72	0.45	0.65	0.79	0.60	0.53
4	Coal + CCS	0.60	0.85	0.55	0.70	0.60	0.58	0.60	0.60	0.85	0.85	0.60	0.62	0.60	0.72	0.45	0.65	0.79	0.60	0.53
5	IGCC	0.60	0.85	0.55	0.70	0.60	0.58	0.60	0.60	0.85	0.85	0.60	0.62	0.60	0.72	0.45	0.65	0.79	0.60	0.53
6	IGCC + CCS	0.60	0.85	0.55	0.70	0.60	0.58	0.60	0.60	0.85	0.85	0.60	0.62	0.60	0.72	0.45	0.65	0.79	0.60	0.53
7	CCGT	0.98	0.85	0.66	0.55	0.66	0.55	0.55	0.55	0.53	0.85	0.71	0.72	0.85	0.55	0.58	0.55	0.30	0.55	0.54
8	CCGT + CCS	0.98	0.85	0.66	0.55	0.66	0.55	0.55	0.55	0.53	0.85	0.71	0.72	0.85	0.55	0.58	0.55	0.30	0.55	0.54
9	Solid Biomass	0.54	0.85	0.50	0.46	0.37	0.50	0.17	0.39	0.69	0.50	0.50	0.50	0.85	0.60	0.50	0.59	0.50	0.50	0.85
10	S Biomass CCS	0.54	0.85	0.50	0.46	0.37	0.50	0.17	0.39	0.69	0.50	0.50	0.50	0.85	0.60	0.50	0.59	0.50	0.50	0.85
11	BIGCC	0.54	0.85	0.50	0.46	0.37	0.50	0.17	0.39	0.69	0.50	0.50	0.50	0.85	0.60	0.50	0.59	0.50	0.50	0.85
12	BIGCC + CCS	0.54	0.85	0.50	0.46	0.37	0.50	0.17	0.39	0.69	0.50	0.50	0.50	0.85	0.60	0.50	0.59	0.50	0.50	0.85
13	Biogas	0.98	0.85	0.66	0.50	0.66	0.50	0.50	0.50	0.53	0.85	0.71	0.72	0.85	0.50	0.58	0.42	0.30	0.50	0.54
14	Biogas + CCS	0.98	0.85	0.66	0.50	0.66	0.50	0.50	0.50	0.53	0.85	0.71	0.72	0.85	0.50	0.58	0.42	0.30	0.50	0.54
15	Tidal	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
16	Large Hydro	0.48	0.85	0.26	0.45	0.28	0.40	0.31	0.57	0.35	0.40	0.33	0.27	0.17	0.29	0.40	0.58	0.40	0.40	0.40
17	Onshore	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
18	Offshore	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
19	Solar PV	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
20	CSP	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
21	Geothermal	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
22	Wave	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
23	Fuel Cells	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
24	CHP	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

Table E.3 cont.

		Rest of Annex I	China	India	Mexico	Brazil	Argentina	Colombia	Rest of Latin America	Korea	Taiwan	Indonesia	ASEAN	OPEC (excl Venezuela)	Rest of world	Ukraine	Saudi Arabia	Nigeria	South Africa	Rest of Africa	Africa OPEC
1	Nuclear	0.78	0.91	0.85	0.82	0.79	0.82	0.85	0.85	0.85	0.95	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
2	Oil	0.30	0.30	0.83	0.70	0.29	0.30	0.30	0.54	0.20	0.36	0.30	0.25	0.51	0.55	0.55	0.55	0.55	0.55	0.55	0.55
3	Coal	0.60	0.58	0.69	0.50	0.85	0.46	0.60	0.43	0.60	0.78	0.74	0.70	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
4	Coal + CCS	0.60	0.58	0.69	0.50	0.85	0.46	0.60	0.43	0.60	0.78	0.74	0.70	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
5	IGCC	0.60	0.58	0.69	0.50	0.85	0.46	0.60	0.43	0.60	0.78	0.74	0.70	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
6	IGCC + CCS	0.60	0.58	0.69	0.50	0.85	0.46	0.60	0.43	0.60	0.78	0.74	0.70	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
7	CCGT	0.55	0.55	0.51	0.52	0.55	0.46	0.55	0.55	0.85	0.36	0.55	0.64	0.49	0.61	0.61	0.61	0.61	0.61	0.61	0.61
8	CCGT + CCS	0.55	0.55	0.51	0.52	0.55	0.46	0.55	0.55	0.85	0.36	0.55	0.64	0.49	0.61	0.61	0.61	0.61	0.61	0.61	0.61
9	Solid Biomass	0.85	0.85	0.50	0.85	0.70	0.85	0.85	0.50	0.85	0.85	0.85	0.50	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
10	S Biomass CCS	0.85	0.85	0.50	0.85	0.70	0.85	0.85	0.50	0.85	0.85	0.85	0.50	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
11	BIGCC	0.85	0.85	0.50	0.85	0.70	0.85	0.85	0.50	0.85	0.85	0.85	0.50	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
12	BIGCC + CCS	0.85	0.85	0.50	0.85	0.70	0.85	0.85	0.50	0.85	0.85	0.85	0.50	0.85	0.83	0.83	0.83	0.83	0.83	0.83	0.83
13	Biogas	0.30	0.50	0.51	0.52	0.50	0.46	0.50	0.36	0.85	0.36	0.50	0.64	0.49	0.61	0.61	0.61	0.61	0.61	0.61	0.61
14	Biogas + CCS	0.30	0.50	0.51	0.52	0.50	0.46	0.50	0.36	0.85	0.36	0.50	0.64	0.49	0.61	0.61	0.61	0.61	0.61	0.61	0.61
15	Tidal	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
16	Large Hydro	0.40	0.40	0.36	0.39	0.54	0.35	0.58	0.64	0.29	0.40	0.27	0.51	0.40	0.38	0.38	0.38	0.38	0.38	0.38	0.38
17	Onshore	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
18	Offshore	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.4	0.4	0.4	0.4	0.4	0.4
19	Solar PV	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
20	CSP	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
21	Geothermal	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
22	Wave	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
23	Fuel Cells	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
24	CHP	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

Table E.4 Energy resources assumptions used to compute cost supply curves in the natural energy resources (NER) module, which an input to FTT:Power.

Resource Name	**Type**	**Dist.**	**Use EJ/y**	**Technical potential**	**Units**
Wind	Flow	Hierarch.	0.72	346	EJ/y
Solar	Flow	Identical	0.04	3384	EJ/y
Hydro	Flow	Hierarch.	12	66	EJ/y
Geotherm.	Flow	Hybrid	0.23	36	EJ/y
Biomass	Flow	Hybrid	51	447	EJ/y
Ocean	Flow	Hierarch.	0.002	23	EJ/y
Oil	Stock	Hierarch.	170	67	10^3EJ
Gas	Stock	Hierarch.	109	46	10^3EJ
Hard Coal	Stock	Hierarch.	139	220	10^3EJ
Soft Coal	Stock	Hierarch.		37	10^3EJ
Uranium	Stock	Hierarch.	30	1.36	10^3EJ
Thorium	Stock	Hierarch.	-	4.68	10^3EJ

Stock/Flow indicates whether resources are renewable flows or stocks. Hierarch./Identical/Hybrid identifies the type of statistical distribution assigned. Use refers to current yearly consumption of these resources.

For the full regional data (for the 59 regions) and equations used to compute cost supply curves in the NER module, as well as an uncertainty assessment, see Mercure and Salas (2012).

Table E.5 Grid flexibility parameter

MRIT = 0.7
The value of the MRIT parameter determines how much flexible technology (Oil, CCGT, Biogas, Hydro) is needed per intermittent technology (Wind, Solar PV, CSP, Wave):