Hochschule für
Wirtschaft und Recht Berlin
Berlin School of Economics and Law

The Implications of Technical Change for Climate Policy and Decarbonization Scenarios

Levi T. Henze
International Economics M.A.
Berlin School of Economics and Law
Student ID: 1841239

January 30, 2023

Master Thesis
Master of Arts in International Economics
First Examiner: Prof. Dr. Eckhard Hein
Second Examiner: M.A. Ryan Woodgate
Date of Submission: January 30, 2022
16989 of 17.000 words*

Abstract

Mitigation scenarios have consistently underestimated the potential of low-carbon energy technologies and recent evidence reaffirms this. I review this evidence and analyze how the most popular integrated assessment models (IAM) treat technical change. Poor technology forecasts are caused by (i) conservative learning rate assumptions (ii) unfounded cost floors, (iii) static operational and capital cost (iv) overall model structure, (v) cost penalties for growth industries, and (v) poor coverage of key technologies. I argue that technical change is a defining feature of the energy transition and must be addressed more thoroughly. With their growing role in guiding policy, IAMs should be calibrated frequently and rigorously. Their results have to be interpreted with care, and model diversity must improve. As to policy, a greater focus on transitional risks, energy rebounds and electrification is needed. The overall economic cost of a net zero economy is likely over-appreciated.

*Excludes lists, tables, figures and references.

Table of Contents

List of Figures

List of Tables

Abbreviations and Units

ABM	Agent-based Model
BECCS	Bioenergy with Carbon Capture and Storage
BEV	Battery Electric Vehicle
CapEx	Capital Expenditure
CBA	Cost-Benefit-Analysis
CCS	Carbon Capture and Storage
CDR	Carbon Dioxide Removal
CES	Constant Elasticity of Substitution
CGE	Computable General Equilibrium
CO_2E	Carbon Dioxide Equivalents
CSP	Concentrated Solar Power
DAC	Direct Air Capture
EROI	Energy Return on Investment
ESM	Energy System Model
ETS	Emissions Trading System
GW	Gigawatt
IAM	Integrated Assessment Model
IEA	International Energy Agency
IIASA	International Institute for Applied Systems Analysis
IMP	Illustrative Mitigation Pathway
IPCC	Intergovernmental Panel on Climate Change
kW	Kilowatt
kWh	Kilowatt Hour
LCOE	Levelized Cost of Energy
MAC	Marginal Abatement Cost Curve
NDC	Nationally Determined Contribution
NET	Negative Emission Technology
OpEx	Operational Expenditure
PJ	Petajoule
PtX	Power-to-X
PV	Solar Photovoltaics
R&D	Research and Development
RES	Renewable Energy Systems
TWh	Terawatt Hour
WACC	Weighted Average Cost of Capital

1 Introduction

What does it cost to abate carbon emissions? This question has been a building block of the economics of climate change from the beginning. The famous DICE model, arguably the first ever Integrated Assessment Model (IAM), aimed at answering how much society should spend on mitigation and adaptation, given the abatement cost, and damages resulting from increasing atmospheric greenhouse gas concentrations. It relies on a typical neoclassical Ramsey growth model in which consumption and investment would be chosen to optimize a long term growth path (Nordhaus 1992). This approach has received both fierce and modest critique. Most critics have argued that the damages from climate change were fundamentally misrepresented, encouraging delayed action (the most prominent example being Dietz and Stern 2015). Some have found the macroeconomic fundamentals of equilibrium analysis to be unsuitable (Pollitt 2019; Grubb et al. 2021b). On the issue of abatement cost disagreement was less of an issue for a long time. The loose consensus was that reducing carbon emissions would prove to be a burden, specifically in less advanced economies. Some partial equilibrium models, now known as Energy System Models (ESMs), tried to get a refined understanding of the economic potential of a decarbonized economy (an early example is Goulder and Schneider 1999). In policy, a crucial hallmark of IAMs, the Marginal Abatement Cost Curve (MAC), which depicts the economic cost over a given amount of abated emissions, became a crucially important policy tool.[1] The approach was questioned for the implication that costs were static and independent, but the underlying idea remained unchanged (Kesicki and Ekins 2012).

Approaching climate change with cost-benefit analysis came under much stronger scrutiny around a decade ago. Climate scientists became increasingly confident about a core pillar of climate analysis: The value of the climate sensitivity, the temperature response to a doubling of carbon dioxide concentrations in the atmosphere. Due to the nature of the data available to us in the near term, the community became confident that its likely range was 2°-4.5°C, with no possibility to produce a narrower range.[2] Weitzman (2009) showed that given this range

[1] In small-scale or analytic IAMs such as DICE, these are approximated by a differentiable function, while more fine-grained, policy oriented work uses a bottom-up approach to evaluate distinct policies.

[2] There are nuances to this metric mainly stemming from our understanding of natural carbon sinks and the dynamics of non-CO_2 greenhouse gases. If human emissions ceded immediately, temperatures would continue rising for about a decade, and then drop to an equilibrium of about 0.3°C lower. While this point seems to be rather clear, the underlying range, the climate sensitivity, remains as a structural uncertainty. In highly simplified terms, this uncertainty can be explained as a signal-to-noise problem. A good summary of this can be found in IPCC AR6 (2022, ch. 3).

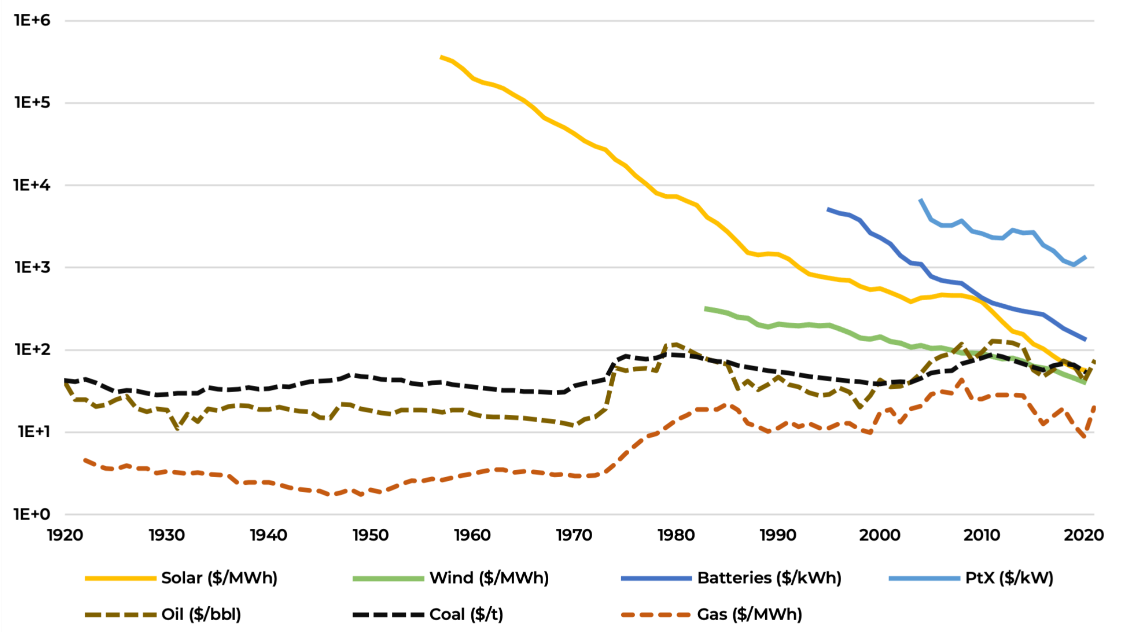
and some very lenient assumptions, the expected damages from climate change become quantitatively unconstrained. With that, the credibility of optimal warming targets, derived from cost-benefit analysis, was shattered. With the Paris Agreement, policy moved instead to carbon budgets aligning with a mutually acceptable peak temperature. The uncertainty remains: For a two third chance of meeting the 1.5°C target, the remaining CO_2 budget is about 300±210 Gt, a range *exceeding* the central estimate (Lamboll et al. 2022). To wit, at constant 2021 annual emissions ($\approx$ 39 Gt CO_2), this puts us about 3 to 13 years away from breaching the primary Paris Agreement target.

Given this degree of urgency, and the irrelevancy of a central tenet of economic analysis, why would we still care about the cost of emission abatement? There are three major reasons. Firstly, one seminal advancement in climate science was the recognition that a stabilized climate would actually require *zero* anthropogenic emissions (Matthews and Caldeira 2008).[3] The *economic* consequence of that is quite radical: Every kind of human activity, give or take, has to switch to a carbon-neutral technology, be compensated through Carbon Dioxide Removal (CDR) or be stopped entirely for temperature stabilization to occur. We may be convinced that emission reductions in large parts of the economy are beneficial even without accounting for damages, but that is certainly not the case for all sectors, communities and regions. Global benefits and damages from emissions are highly unequally distributed (Oswald et al. 2021). Trillions of dollars worth of fossil fuel assets incompatible with the Paris Agreement targets have to be stranded (Semieniuk et al. 2022). And with fossil fuel demand likely peaking this decade (IEA 2022), there are transitory costs to consider, resulting from the rapid down scaling of a deprecated energy system (Grubert and Hastings-Simon 2022).

Secondly, since decarbonization essentially means the complete replacement of the fossil capital stock, the macroeconomic effects could be sizable (Pisani-Ferry 2021). There is good evidence that, beside regional disparities, inflationary effects will be small and employment and growth effects positive (Pollitt et al. 2017). However such an assessment would look quite differently without the success of Renewable Energy Systems (RES) in recent years (see figure 1). Widely acknowledged today, this is the outcome of industrial policy mainly in Germany and China (IRENA 2021). But this is not what many IAMs scenarios would have predicted relatively recently (Way et al. 2022). From a macroeconomic perspective we would want to know

[3] There is a large uncertainty to that, but this goes both ways: Temperature stabilization might need carbon removal *in excess* of gross emissions or leave room for a residual, with a likely range of about ±4 Gt Carbon Dioxide Equivalents (CO_2E) (Allen et al. 2022).

Figure 1: Historical Cost of Energy Supply Technologies



Cost curves of energy technologies in historical comparison, log scale. Wind and Solar: global average Levelized Cost of Energy (LCOE) (IRENA), Batteries: cell costs (IEA/BNEF), PtX: capacity cost of proton-exchange membrane electrolyzer (IEA), Oil: Brent Crude (FRED), Gas: US wellhead price (EIA), Coal: US average bituminous (EIA). Prices are not comparable in terms of useful energy, but electricity generation from fossil fuels has not notably improved over time. Source: Way et al. (2022), own calculations.

either way. Naturally, the cost of useful energy plays a large role in that. Deployment of RES has also been the strongest driver of decarbonization. This trend is very pronounced in countries where domestic emissions have peaked and efforts are supported by deployment policy (Le Quéré et al. 2019). A reduction in energy consumption has contributed, but it is mostly the declining share of fossil energy that has supported decoupling trends (see table 1). But the world economy also emits well above 12 Gt CO_2E of *additional* greenhouse gases from non energy use related activity, such as agriculture and land use (UNEP 2022). Parts of the transition are already declared sensible business choices and development pathways (Wolf 2022; Burn-Murdoch 2022). But the tail end of the abatement challenge still rests on far shakier grounds.

Lastly, the specific choices as to which technology to use and at what fraction, is going to depend on their cost. Finding the one cheapest path to zero emissions is decidedly a futile exercise, but that does not mean cost (and cost potential) are irrelevant. In order to describe plausible decarbonization scenarios and their outcomes, we have to make well-founded assumptions about technology cost, material availability, potential scalability, labor requirements and so forth. Existing models have had a bad track record with that. To choose a particularly puzzling example, one core scenario utilized in IPCC SR15 (2018) featured carbon removal on the order of 20 Gt CO_2 per year, which amounts to roughly half of current annual emissions. We would rightly doubt the physical possibility and cost-effectiveness of that. Given that national climate policy is strongly influenced by the work of the Intergovernmental Panel on Climate Change (IPCC) and the International Energy Agency (IEA), scrutiny is warranted about whether leading scenarios and models give useful guidance.

Table 1: Decomposition of Emission Reductions in Post Peak Countries

Percent reduction due to...	energy use	fossil share	fossil efficiency	fossil emission intensity	trade
Maximum	190	133	39	38	8
75th Percentile	56	73	22	4	0
Median	36	47	12	0	-3
25th Percentile	18	36	-14	-7	-14
Minimum	5	25	-85	-55	-78

Drivers of observed emissions reduction (2005-2015), as percentage of cumulative reduction, in a sample of 18 developed countries where emissions have peaked. Country-wise, the first four columns add up to 100 percent, trade is depicted as a leakage control. Efficiency reference is primary to final energy. Sample includes Austria, Belgium, Bulgaria, Croatia, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Netherlands, Portugal, Romania, Spain, Sweden, United Kingdom and USA. Source: Le Quéré et al. (2019).

All three problems have a common focal point: The cost of technology. Arguably then, a shift away from cost optimization to cost-effectiveness, as it is often framed, makes a closer look into the dynamics of technology deployment even more important.[4] But can we predict how costs will develop over time? All too conservative estimates have consistently resulted in pessimistic outlooks. But are these cost reductions a ubiquitous phenomenon, and would they extend to a decarbonized economy as a whole? Some observers are not convinced that technical change is something we can rely on (Pisani-Ferry 2021). Modelers might argue that self-defeating prophecies are a useful tool in policy advocacy and hence conservative assumptions about technical progress are helping. But of course this can go both ways and stall progress instead.

This paper is an attempt to take stock of the situation. How are mitigation scenarios and models addressing technical change? How have they fared so far and how can we improve them? Two answer these questions, the rest of the paper proceeds as follows. After discussing some methodological questions in section 2, section 3 will give a short review of how mitigation scenarios performed in meta-studies of technology cost and deployment trends. This is the evidence that necessitates a closer look into IAMs assumptions. Then, section 4 discusses aspects

[4]See IPCC AR6 (2021, ch. 1.7) for context on how the 'effectiveness' framing came about.

of technical change in mitigation modeling, in turn: Formal implementation (4.1), calibration (4.2), performance metrics models use (4.3), interaction with broader model design (4.4) and finally technology representation outside the energy sector (4.5). Section 5 will conclude and provide avenues for further research.

2 Remarks on Methodology

The questions set out in section 1 warrant some strategic remarks on how to approach them. To start, I focus specifically on *global* assessments since (i) national studies often use similar or derivative models, (ii) global assessments inform nation or region-specific research and (iii) the choice of technology representation follows very different theoretical considerations in a national context. The latter are mainly due to market size relative to national policy scope, spillovers, and regional variations (see these issues explored in Crassous et al. 2006). National incentive structures arising from technical change, are expressly not the focus here. The next issue concerns the selection of scenarios. In following the global scope, it makes sense to probe recent work of the IPCC and the IEA. These are arguably the two most influential institutions when it comes to climate policy on a global level. Assessments of both institutions will feature in section 3, a review of comprehensive scenario assessments. The focus on models, as opposed to scenarios, warrants justification. It is apparent from scenario assessments that the underlying models themselves must have introduced certain biases into global mitigation scenarios. This is why the rest of the discussion (section 4) focuses on model structure and implementation. It is essentially what scenario assessments recommend at the current state of research.

It turns out that this makes progress on finding definitive answers very hard. For a start, due to largely obscured modeling decisions it is not possible to discuss the work of the IEA in more detail. This only leaves the IPCC, specifically the most recent assessment report (IPCC AR6 2021). It's role in relation to climate science is twofold: It builds a comprehensive and regular review of the progress in recent literature and systematizes it. It also guides further research: In mitigation, for example, by highlighting deficiencies in models and scenarios, or by constructing guideline scenarios such as the Illustrative Mitigation Pathways (IMPs) (explained in IPCC AR6 2021, ch. 1). These are also used to classify research, and draw out trends that correspond to one another across simulations.

With the current assessment cycle the IPCC author community has undergone efforts to increase transparency of scenario design. The scenario database allows for comparison between thousands of model runs and studies (IIASA 2022). A useful way to approach this whole body of work is laid out by Koomey et al. (2019), who suggest decomposing macro trends in mitigation scenarios in a layered fashion, as an expansion of the Kaya decomposition. In theory this would be a viable and effective approach for the question at hand, laying out which model-specific assumptions drive technology adoption. Brockway et al. (2021) use it, for example, to identify a potential underestimation of the macroeconomic rebound across scenarios. But in the case of technology deployment the impacts on macro trends of the economy are not as straightforward. For example, a scenario deploying large amounts of carbon removal could achieve similar carbon intensity reductions as another based on renewable energy, but would result in vastly different technology choices. Yet, some issues discussed in section 4 could indeed be analyzed in a data-driven approach using the scenario database, and this paper can be seen as a foundation for such research.

If scenario assessments point distinctly at models as a "culprit" but there is no clear idea as to how to approach model output systematically, this leaves the models themselves to explore. This will be the core of this paper. But model documentation is still a bottleneck even though there has been a concerted effort to increase the transparency of IAMs modeling (IAMC 2022). To the extent possible on that basis, model specifications will be discussed in section 4.

This leaves as question which models give a representative sense of current mitigation assessments. Table 2 shows the simulation shares of the ten most used models for IPCC AR6 (2021) together with their type and scope. In the past, authors have often classified models along different dimensions or according to their theoretical background (Crassous et al. 2006; Farmer et al. 2015; Dafermos and Nikolaidi 2022). At the current point in time these descriptions are largely unhelpful. Due to their advanced, modular design, models can change fundamental aspects of their solution space, for example being able to run in general and partial equilibrium mode. Hence today, the model type can mostly be interpreted as a guiding principle or modeler philosophy. This may have some bearing on model outcomes regarding technology, briefly discussed in section 4.4, but that is rather speculative. Finally, the documentation column shows the sources and their respective dates for anything I will discuss related to specific model assumptions. Notably, this information is up to a decade old and studies relying on these models may make modifications. The author has gone to some length to access the most recent information, but this does not mean studies featured in IPCC assessments have used these models without major alterations. This is why I take a wider view, discussing a breadth of *possible* issues.

Choosing a set of models to discuss is made easy by lack of diversity. For Paris compliant

Table 2: Integrated Assessment Models for the IPCC's Sixth Assessment Report

Model	Type	Documentation	Regional Scope	Percent share modeled	
				<2°C	all scenarios
REMIND	CGE, perfect foresight	Luderer et al. (2015)	Global	26	17
MESSAGE	ESM, limited foresight	Krey et al. (2020)	Global	23	16
IMAGE	system-dynamic, myopic	Stehfest et al. (2014)	Global	10	8
POLES	ESM, myopic	Keramidas et al. (2017)	Global	9	7
WITCH	CGE, perfect foresight	Drouet et al. (2019)	Global	8	9
COFFEE	CGE, perfect foresight	–	Brazil	5	4
GCAM	ESM, myopic	Bond-Lamberty et al. (2022)	Global	5	8
AIM	CGE, perfect foresight	–	Japan, China	5	9
GEM-E3	CGE, myopic	Capros et al. (2013)	Global	4	3
TIAM	ESM, perfect foresight	–	UK	4	4
other		–	–	1	14

The ten most commonly used IAMs in IPCC AR6 (2022). Leftmost columns show (i) the share of modeled scenarios that are compatible with the Paris Agreement, with a 50% chance of staying below 2°C (n=700) and (ii) of all scenarios reporting final emissions (n=1861). Model type (Computable General Equilibrium (CGE),ESM, system-dynamic) refer to the modeling philosophy regarding macro trends, since most models can run in either a partial or general equilibrium today.
Sources: IAMC (2022) and Krey et al. (2019) and own calculations based on IIASA (2022).

scenarios, 76 percent of the simulations are driven by five models, for *all* 1800 scenarios which report emission data that figure is 58 percent (last two columns in table 2). I simply focus on the ten most used ones, minus those with a regional scope.[5] It is still hard to show anything conclusive about the consequences model assumptions. Sognnaes et al. (2021) and Krey et al. (2019) use an inter-model comparison approach to probe model assumptions with respect to CDR utilization.[6] Something similar has been done in the past when many models introduced technical change into their structure (Crassous et al. 2006). Alas, such an approach exceeds the scope of this work. Before it would yield anything conclusive, the possible roots of biases present in the existing model set would have to be explored anyway.

Due to these issues of transparency and scope, the discussion to follow has to remain largely exploratory in nature. Identifying critical assumptions quickly becomes an archeological exercise, and documentation is often not comprehensive enough.[7] Some leads can also prove

[5] If we consider a broad model base from diverse backgrounds to be of value, this lack of diversity is less than optimal.

[6] The term *Carbon Dioxide Removal* is not used fully consistently in the literature. I refer to the definition of IPCC AR6 (2021, Annex I), where it is taken to mean any kind of durable storage of atmospheric CO_2, including anthropogenic enhancement of carbon uptake, such as reforestation or soil sequestration.

[7] The most widely used models, such as REMIND, IMAGE, MESSAGE or WITCH are open source platforms today, but documentation efforts are not always up-to-date.

inconclusive, since without running simulations there is no way of knowing whether certain constraints are binding or not. It should still be kept in mind that the matters to be discussed effectively cover thousands of studies, because these models are used so extensively all around the world. Hence, short of developing new models from scratch (Dafermos and Nikolaidi 2022; Hoekstra et al. 2017), a black box approach could identify common biases (Koomey et al. 2019) and sensitivity studies could probe model constructions (Sognnaes et al. 2021). This work intends to map out a foundation for all three of these areas of research.

A final remark regarding terminology: I refer to the sample of models discussed here as *large-scale IAMs* since differentiating them from analytical models is the only sensible ad-hoc distinction one can draw at this stage. Theoretical background and development history almost certainly plays a role (see section 4.4), but given their flexible, platform-like nature, outcomes depend on a host of other assumption sets as well.

3 Technology in Ex-Post Scenario Evaluation

The *physical* potential of low carbon energy sources, only considering yield, environmental and land-use constraints, never was in much doubt.[8] Even two decades back, the IPCC estimated that the potential Solar Photovoltaics (PV) alone exceeding primary energy demand projections of this century (Creutzig et al. 2017). In contrast to that, the underestimation of actual deployment successes has been consistent in research, industry and policy work for several decades now. There is a small but comprehensive number of meta-studies on past scenario performance that demonstrates this. The initial focus was put on PV deployment trends, arguably the most obvious one, but biases have recently been shown to extend into cost predictions as well and span a wider range of technologies critical to decarbonizing the energy supply.

The first suspicions of systematic biases were raised surrounding the over-reliance on bioenergy. Creutzig et al. (2017) show that assuming endogenous technical progress, 50% of electricity could come from PV in a cost-efficient energy system. The study analyzes hundreds of scenarios including from the IEA, the IPCC and Greenpeace. Across the board, these were inconsistent with historic deployment trends of PV. With predictions less than half of actual deployment in 2015, IPCC scenarios for the Fifth Assessment Report were already hopelessly out of date upon final publication. Generous policy support for RES, increased costs among

[8]This is following the IPCC definition of 'technical potential' IPCC AR6 (2021, Annex I), which intends to capture *deployment* limitations as opposed to production limitations.

competing technologies (mainly Carbon Capture and Storage (CCS) and nuclear power) and "atypical technical advancement" were seen as likely reasons (IPCC SR15 2018). It seems plausible, in hindsight, that the physical similarity of bio-energy to fossil fuels, and its apparent advantages in terms of transport and storage, has contributed to the relative weight researchers put on it. Net zero – the idea that climate stabilization demanded zero anthropogenic emissions – was also relatively new (Allen et al. 2022). In a world where net zero would *not* be a necessity, the land-use requirements of large scale bio-energy production certainly are less of a problem.[9] It is worth mentioning that today's installed PV capacity of about 760 GW only just lies in the upper range prediction of Creutzig et al. (2017). For IPCC AR6 (2021) a major focus was to correct for the over-reliance on CCS, but that did not exactly solve the pessimistic assessments regarding RES.

To understand this systematic pessimism a little better, more recent scenario assessments focus directly on cost forecasts, as opposed to deployment. Early climate economics literature frequently discussed the possibility that solar and wind could become cheap, replacing more expensive fossil sources gradually (Crassous et al. 2006; Nordhaus 2013). Today, there is ample evidence that large parts of the energy supply can be economically replaced by RES. Studies differ as to the extent of that, but it is directly visible from cost comparisons[10] (IRENA 2021), it has been shown system-wide, for example for the European power sector, with and without factoring in the Emissions Trading System (ETS) (Victoria et al. 2020; Rosslowe and Cremona 2022), for a two-third decarbonization of power in China (He et al. 2020) and also globally (Bogdanov et al. 2021). Of course such studies are forecasts and have to rely on ESMs or IAMs themselves. In other words, the question of forecasting accuracy and bias remain, even though the outlook for RES has already changed dramatically.

Evidence of underestimation of PV potential is unequivocal, but model features and scenario design are hard to entangle. Jaxa-Rozen and Trutnevyte (2021) used a statistical learning methodology on a large sample of IPCC, non-IPCC and gray literature mitigation studies to identify what drives potential estimation in terms of deployment. Some significant influencing factors are quite obvious: The publishing institution, climate policy assumptions and

[9] A good summary of the problems associated with bio-energy can be found in (IPCC AR6 2022, ch.7). When accounting for food security, the global potential in 2050 is estimated to be less than half of today's primary energy demand, which does not account for the energy needs in storing the carbon produced at that scale. Impacts on other ecological dimensions and a potential to worsen anthropogenic emissions from land use are significant barriers as well.

[10] See section 4.3 for more detail

publication date influence PV deployment strongly. However, much variation is indeed driven by model construction. Analytical models, as opposed to the large-scale IAMs featuring in IPCC scenarios, estimate stronger solar power deployment. Whether technical change is endogenous or not plays only a minor role (for *deployment*, importantly), and so do foresight (myopic/perfect) and model closure (partial or general equilibrium). Various types of hard and soft deployment constraints are the most important factors.[11] Jaxa-Rozen and Trutnevyte (2021) propose stronger model diversity to account for this. Their sample was taken from the fifth assessment report, but so far such proposals have not changed the dominance of certain models (see table 2).

Victoria et al. (2021) look more closely at the technology potential of PV and find dated cost assumptions in most IAMs. Good spatial and temporal resolution seem to favor PV, the former because it reduces rivaling uses compared to more stylized "frictional" representations and the latter for reducing various matching potentials between intermittent sources. Increasing resolution, especially in the temporal dimension, also faces computational trade-offs.[12] On the cost side, they point to a steady decrease of module prices, but they do not probe models with respect to that. They point to significant potential due to the majority of the world's population living in areas where solar flux is not seasonal. Of course such a single technology perspective has tight limits and can overlook crucial aspects of decarbonization or growth potential. In the case of PV, significant resources need to be expended for transmission, storage, sectoral coupling (electrification of power demand) and grid stability (Victoria et al. 2021; Creutzig et al. 2017; Parzen et al. 2022). Yet, quite conclusively, models underestimated deployment *and* cost reductions of PV.

A much smaller sample of 22 studies (IEA and the most prominent IAMs) is probed by Xiao et al. (2021), focusing explicitly on cost assumptions, but this time for wind and solar energy. Interestingly, they not only look at investment cost or Capital Expenditure (CapEx) but also compare LCOE, a standard per-energy metric for power supply technology, with current auction results.[13] For large scale supply such as Concentrated Solar Power (CSP), offshore wind and utility scale PV, they find that all but one of the studies examined predict lower cost in 2050 than what auction results were in 2019. This striking mismatch is despite the fact that in

[11] They differentiate between hard constraints as fixed growth limits and soft constraints as adjustment costs.

[12] The modeling strategy of the IEA is a good example. In the past, their model used hourly resolution on a global scale, but in order to make that feasible, explicitly refrained from endogenous technology representation (IEA 2021).

[13] The question of cost metrics is discussed in section 4.3 and LCOE are explained in the appendix.

terms of CapEx, the underlying studies differ in outcomes by a factor of five. Even the most optimistic model assumptions result in bad predictions. Also, modelers accounted for technical progress affecting the investment cost but neglected other, more important dimensions along which technologies change. A common suggestion to improve data and update model assumptions regularly is echoed here (Xiao et al. 2021).

A very recent study by Way et al. (2022) expands on previous efforts and garnered widespread attention in climate policy debates (Ives 2021). They compare IPCC scenarios in the sixth assessment report to a simple but complete representation of a decarbonized power system comprised of solar, wind, battery storage and Power-to-X (PtX).[14] The analysis employs non-deterministic forecasting for technology, based on previous work that aimed to improve the empirical foundation necessary for mitigation modelling (Farmer and Lafond 2016; Nagy et al. 2013; Lafond et al. 2020). The scenario assessment reaffirms what others have found before: Even the most recent and most optimistic IAM predictions for IPCC AR6 (2021) assume noticeable, structural breaks with historical technology trends. Crucially the trends analyzed here are not in terms of deployment or investment cost, as in the literature discussed before, but in terms of LCOE and as a function of cumulative deployment. In a fast transition scenario, which assumes historical growth rates for the aforementioned technologies, the overall system *saves* several percentage points worth of global GDP over the next decades. This is compared to a no-transition baseline scenario, mirroring common assumptions in IAM modelling. Though simplified, the model of Way et al. (2022) demonstrates conclusively that existing IAM vastly overestimate the costs of key low-carbon technologies.

The common threads in the literature discussed above are twofold. The first one is that ex-post scenario evaluations reveal a persistent, widespread and surprisingly large underestimation of the economic potential of low emission energy supply technologies. The second one is that so far nothing seems to indicate that specific assumptions of single studies are to blame, but that instead scenarios seem to inherit relevant characteristics from the underlying models. To the author's knowledge no one so far has suggested or demonstrated that any single model fares significantly better than others. None of the popular models stand up to scrutiny with respect to technology potential and Way et al. (2022), IRENA (2021), and Victoria et al. (2021) make the convincing case that there are no obvious reasons why historical cost trends should markedly slow down. This means that the models urgently need adjustments

[14] PtX refers to a variety of solutions aimed at converting electricity to some other form of energy. For the power system itself, it mainly serves as storage and usually involves hydrogen from electrolysis.

How the IPCC frames technical progress

Before discussing further details of model assumptions it is useful to take a look at how the IPCC currently views these issues. The 6th Assessment Report features an entire chapter on technical change and technology representation in mitigation scenarios (ch. 16 IPCC AR6 2021). The report takes a rather broad sweep through the issues related to technical change, which corresponds to the assessment that technical change needs deeper study and is a major source of uncertainty for climate policy. However, it summarizes some general characteristics of technical trends that seem to be fairly well established across the literature (see Koh and Magee 2008; Grubb et al. 2021a; Grafström and Poudineh 2021; Wilson 2012; Rodrik 2014). (i) Technologies move through distinct stages of development, deployment, diffusion and saturation. (ii) Technical progress is closely linked to economies of scale, and technologies that are suited for mass production tend to realize steeper cost reductions. (iii) There are issues arising from the problem of system boundaries and proper metrics to assess costs, investments and improvements over time. (iv) Technical change requires a broad range of policy considerations, introducing spill-over, but also first-mover advantage. (v) Modelling continues to have considerable difficulties in properly accounting for technical change. For current climate scenarios, (vi) IAMs "tend to underestimate innovation in energy supply but overestimate the contributions by energy efficiency" (IPCC AR6 2021, ch. 16, p. 27).

Now, a central goal of the IPCC reports is to provide a comprehensive summary of the current scientific view on issues related to climate change. Evidently, it is not a lack of awareness that lead to the persistent underestimates outlined in this section. This reaffirms the suspicion that the socioeconomic assumptions made in models are off and need revision. I will take a broad approach to that and discuss the aspects that drive *cost* projections first (sections 4.1 to 4.3). Way et al. (2022) and Victoria et al. (2021) argue that *deployment* forecasts have improved, but not conclusively. This is why sections 4.4 and 4.5 also discuss possible barriers to deployment in scenarios.

4 Technical Change in Integrated Assessment Models

The way technology does or does not progress has always been a core question in environmental economics. The lack of a feasible technological alternative to coal power drove the worry of Jevons (1865) that England's industrial revolution may be short-lived. Georgescu-Roegen (1981), one of the forebears of ecological economics, emphasized that technology ultimately cannot surmount the physical limits of the earth. Stiglitz (1980) responded with optimism, pointing to human ingenuity always finding new ways to make use of earth's endowments. It may be the long-term perspective inherent to environmental questions that has repeatedly divided positions into static and dynamic views of technology.

Oddly, early climate economics sits firmly in the static camp. The predecessor of modern IAMs provided a link between neoclassical growth theory and the physical understanding of our atmosphere (Nordhaus 1992), but it did so with the baked-in assumption that carbon is immutably of value to the economy. Emission abatement would necessarily incur a cost on society. As we have seen in section 3 that is a very odd assumption with mountains of evidence contradicting it.[15] The technology set available to us changed drastically over time and was heavily influenced by mitigation policy.

Technical change: An analytical frame

How does that change the way we think about the cost of abatement? Attempts to acknowledge technical change in an analytical, stylized way are surprisingly recent. Here, I use Grubb et al. (2021b) and Grubb et al. (2022) to show how the temporal dynamics that result from endogenous technical change can generally be understood. This formalizes some basic conclusions that are worth highlighting before looking into more complex model setups. A simple DICE-like abatement cost curve, that denotes annual welfare cost of emission abatement, such as in (Nordhaus 1992) serves as a comparison.

$$\text{DICE abatement cost: } C_t = a(t)^{\theta} \text{ with } a(t) = E_{ref} - E(t) \tag{1}$$

In DICE, annual abatement cost C_t depend on the amount of abated emissions $a(t)$, which are represented as a deviation of annual emissions $E(t)$ from a reference E_{ref}. An exponent θ denotes that the larger the fraction of abated emissions, the more costly *marginal* abatement (C_t') becomes. To account for a change in technology, θ is re-calibrated, which is what has been done in DICE model updates (Grubb et al. 2022). Note that in equation (1) abatement incurs constant cost over time: Unless the fraction of emission reduction changes, annual costs stay

[15] There is the question about whether cost-efficient abatement should be included in the baseline when constructing mitigation scenarios (Nordhaus 2013). But this is quite literally an academic debate and is almost never done well in practice (See for example IEA (2021) and discussion in Way et al. (2022)).

constant.[16] This would be perfectly adequate if we believed that *all* progress in technology is exogenous: There is no need to anticipate cost reductions from policy, or demand. In fact, quite the opposite: *Waiting* for progress – a change either in θ or E_{ref} – would be rational. The implication of *endogenous* change is the exact opposite: Some fraction of the abatement cost will recede. A simple way to approximate this behavior is a "pliable" abatement cost curve.

$$\text{Pliable abatement cost: } C_t = (1-p)\, a(t)^{\theta} + p\hat{t}a'(t)^{\theta} \tag{2}$$

In this representation, a share $p \in [1,0]$ of the abatement cost is dependent on the derivative $a'(t)$. A year-on-year *increase* in abatement is subject to this fraction of the abatement cost. A second parameter, $\hat{t}$ represents the time frame it takes for learning processes to materialize, such that a shock-like abatement is still more costly than a managed and gradual one.[17] Initially, abatement costs are high and radical policy shifts do not pay off. But together with the temporal, pliable element the economy follows path dependency in both directions. A given level of ambition incurs cost now, but makes decarbonization cheaper after implementing mitigation policy. Depending on p, a large or a small fraction of the abatement cost would recede over time.

This also highlights a potential *downside* of technical progress. If clean energy becomes cheaper, not only do abatement costs fall but energy demand may grow more strongly. This macroeconomic rebound would offset some decarbonization progress – scenario realism in *that* regard is also disputed (see especially Brockway et al. (2021)). But this is also, as Way et al. (2022) call it, "a good problem": It enables clean leapfrog growth for developing countries and emerging markets, that historically followed less carbon intensive growth paths than industrial nations (Burn-Murdoch 2022). "Green" rebounds also typically do not follow the same carbon intensity as the overall economy, as instead, clean technologies crowd in greener production along their supply chain (Rosenbaum 2019). This effect is strengthened, if inputs are allowed to progress themselves, counteracting fossil technology lock-in.[18]

[16] This representation dos not recognize market distortions, distributional or capacity effects, financial constraints and much more. Some of these omissions have more deeply rooted consequences that cannot be captured by a simple abatement cost curve (see (Pollitt 2019; Grubb et al. 2013; Dafermos and Nikolaidi 2022)). Both equations (1) and (2) are further simplified here for demonstration purposes.

[17] Ironically, this "inertia" characteristic is what may be a separate cause for slow adoption of new technologies in IAMs (see section 4.4).

[18] Note that this refers to the emission rebound, not the overall energy rebound or macroeconomic rebound. A clarifying discussion can be found in Lange et al. (2021).

4.1 Formalization of Technical Change

In today's large-scale IAM, abatement policy is also subject to cost constraints, but reduced-form abatement cost such as in equations (1) and (2) are replaced by a fine-grained bottom-up representation of the energy system and other sectors.[19] In other words, a microeconomic representation of technology is needed for these models. I do not intend a comprehensive discussion of the economics of technology, technical change and its deep implications in economic theory, but it is useful to give some context.

What constitutes technology is ultimately determined by a researcher's perspective: We can, for example, differentiate a procedural and an input-output view (Dosi and Nelson 2010). The former would be interested in the various interactions comprising its production or its use, what we could call *routines*. The latter would be interested in defining precisely the outcome of an equally precisely defined set of requirements, what we could call a *recipe*. The term 'technology' is also historically closely linked to knowledge. Of course, economic theory is much more comfortable with the *recipe* concept: It is often interested in functional outcomes (Hoekstra et al. 2017). Naturally, modeling does not really have a meaningful choice here: Models need to imply a recipe, linking inputs and outputs. We can classify technical *change* in terms of its functional impact, loosely following Junginger and Louwen (2020).

(i) **Through coincidence, or disruptive innovation:** By definition this kind of progress cannot be anticipated, which implies that it also cannot be modeled. Sometimes, a backstop technology serves as a stand-in for such ill-defined technical possibilities (Nordhaus et al. 1973).[20] Note that model omission should not be equated with irrelevance and separate empirical literature on this topic is very broad (Grubb et al. 2021a).

(ii) **Through time, or learning-by-waiting:** This represents purely exogenous technical change, and corresponds to equation (1). On a microeconomic level it is sometimes also referred to as *Moore's Law*. Moore found a doubling rate over time of components in integrated circuitry (Moore 1965). Such regularity is observed in many industries (Farmer and Lafond 2016).[21]

[19] A few economists disliked this direction of structural realism, that intended to capture systemic interactions which drive macro-dynamics. The unavoidable lack of transparency is certainly a drawback, but it seems the alternative, proposed for example by (Pindyck 2017), has not much to offer in return.
[20] WITCH features an explicit backstop for nuclear power generation, for example.
[21] The term "learning-by-waiting" is never used in the literature, but it is an apt description of what it implies.

(iii) **Through research, or learning-by-searching:** This is a form endogenous technical change whereby research is driving progress. It originally found its way into IAMs because of the similarities between their neoclassical macroeconomic core and the work by Romer (1989). Computational and analytical issues made it hard to go beyond that (Crassous et al. 2006). In large-scale IAM the same concept is also applied to single technologies.

(iv) **Through deployment, or learning-by-doing:** Fully endogenous technical change assumes that technologies improve simply by producing more of them. This has deep roots in the history of economic thought (Arrow 1962; Wright 1936). There have been many proposals in the past on how it works, some of which are discussed below. In technology-specific contexts it is also referred to as *Wright's Law*.

The latter three have all found implementation in large-scale IAMs to varying degree. But it is learning-by-doing that will be the main focus of the discussion, since this is overwhelmingly suggested by the evidence discussed in section 3 and evidence on single technologies (Farmer and Lafond 2016). A brief formal discussion will provide some clarity. Again, the primary focus of this section is not a *causal* theory of technology, although the mechanics of statistical aggregation and the endogenous nature of economic processes have indeed been proposed to drive Wright's Law (Yelle 1979; Sahal 1979; Farmer and Lafond 2016).

Assuming we have unbiased metrics for technology cost and performance (say, the market price and energy output), formalization and estimation of (i) to (iii) as well as combinations of these are straightforward. Starting with exogenous progress, Moore (1965) expressed technical progress as a doubling rate over time.

$$\text{Moore's Law: } C_t = C_0 e^{-\alpha t} \tag{3}$$

Technology cost C_t are a simple exponential of time t with a constant α, scaling the initial cost C_0. Solving for the time it takes to cut the cost in half yields the doubling time:

$$\text{Doubling time: } T_d = \frac{\ln 2}{\alpha}, \text{ with } \alpha = \ln\left(1 + \frac{r}{100}\right) \tag{4}$$

This way, the Moore's law exponent can be expressed in terms of a growth rate r. This relation is inexorable: Nothing can influence technical progress and nothing can prevent it. Empirically, it is nothing more than an exponential trend projection (Arrow (1962) referred to it as "a confession of ignorance"). It cannot satisfy from a theoretical standpoint – particularly if ex-ante policy evaluation, such as is the case with IAMs, is the focal point. But models do use it, or even assum constant cost (Grubb et al. 2021b).

What would drive progress other than the mere passing of time? The idea that experience plays a role in production cost is immediately intuitive and played a major role in post-war economics (Hogan et al. 2020). Its empirical validation is also older than Moore's work and goes back to Wright (1936). Wright discovered a regularity in the manufacturing cost of airplanes, explained by cumulative units produced.

$$\text{Wright's Law: } C_y = C_0 y_\Sigma^{-\beta} \tag{5}$$

Here, technology cost C_y only depend on the cumulative units produced, y_Σ. Interpreted narrowly again, it represents pure learning-by-doing, where y_Σ^β is the cost scaling factor of the next unit.[22] The resulting functional relationship is often referred to as the *learning curve*. For the simple Wright's Law we can define a constant learning rate, which is the cost reduction a doubling in cumulative production yields.

$$\text{Learning rate: } l = 1 - 2^\beta \tag{6}$$

Again, the basic notion of Wright's Law has found overwhelming evidence across many industries. There are many variants to it and different approaches to causal explanations. Arrow (1962) interpreted "learning" very literally, and thought of it as the improvement of workers' skills. Goddard (1982) interpreted it as pure economies of scale and specified it over annual instead of cumulative production. Sahal (1979) gives statistical intuition to how it arises from a process with multiple inputs. Section 4.2 aims at discussing some empirical issues surrounding it, but it describes technological progress remarkably well.

That is, up to a certain point: Particularly the front and tail end of the log-linear learning curve seemed not to match complete adoption cycles very well. Some studies assumed concavity on the left side, meaning slower cost reductions at the beginning (Yeh and Rubin 2012). Similarly, the tail end seemed to be in doubt in long-term studies, suggesting a decreasing learning rate and asymptotic behavior of cost (Klepper and Graddy 1990; Yelle 1979).[23] One intuition

[22] For the purpose of clarity I denote the negative exponent sign to indicate falling cost.

[23] The combination of equations steepening and flattening leads to an S-shaped deployment curve that is now widely regarded as the typical adoption cycle for scalable technologies (**way2021**; Crassous et al. 2006). The resulting specification is a logistic function that is log-symmetric, something that also finds widespread use in

for these floor costs is that some elements of a production process may benefit from learning, while others do not, the latter taking up an increasing share of unit cost. Curiously, IAMs only utilize a floored variant – assuming a steepening by default would have improved the anticipation of PV cost reductions section 4.2. An asymptotic variant of Wright's Law is sometimes referred to as a DeJong specification.

$$\text{DeJong Model: } C_y = C_0 \left(\theta + (1-\theta)\, y_\Sigma^{\beta}\right) \tag{7}$$

It introduces θ as a factor of compressibility, representing a fraction of the production process that can be improved, while the rest cannot. This introduces a floor cost, at which a technology will not improve further. Some models approximate this behavior by using a step function, meaning a Wright's law specification with a floor cost at a threshold.

$$\text{Floor cost: } C_f = \lim_{y_\Sigma \to \infty} C_y = \theta C_0 \tag{8}$$

Lastly, learning-by-searching is often implemented in conjunction with learning-by-doing to better represent nascent energy technologies. In the early stages of development and deployment, markets are often limited and funding is always split between expansion of capacity and Research and Development (R&D) (Jamasb 2007). This led to the two-factor learning curve, proposed for energy models by Kouvaritakis et al. (2000).

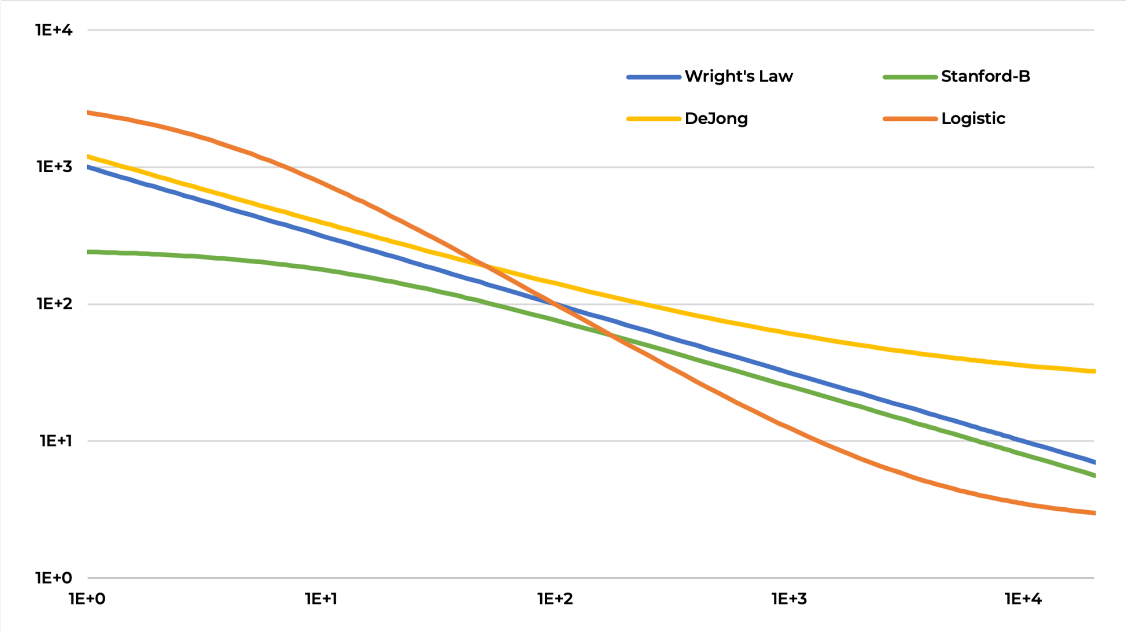
$$\text{Two-factor learning: } C_y = C_0 Y_\Sigma^{-\beta} R_\Sigma^{-\delta} \tag{9}$$

This is also commonly dubbed a learning-by-searching model. In this case, R_Σ denotes R&D spending at time t. In the context of mitigation modeling this typically means policy allocation of research funding, but it has also been used to investigate deployment subsidies and crowding-in (Wene 2008).

For intuition, figure 2 shows a stylized log-log representation of Wright's Law (equation (5)), a steepening learning curve (Stanford-B model), the DeJong model (equation (7)) and a logistic deployment path. Logistics and Stanford-B representations are not used in the model sample so omit a formal presentation. However, logistics are a common feature when it comes to resource depletion (Luderer et al. 2015; Keramidas et al. 2017; Dafermos and Nikolaidi 2022).

ecological models.

Figure 2: Learning Curve Specifications



Stylized log-log representation of different learning functions. DeJong and Stanford-B behave asymptotic to Wright's Law, with the former approaching a cost floor and the latter steepening. Wright's Law converges to zero, which is why it is often considered unrealistic. The logistic combines the two asymptotic variants to match a full, stylized adoption cycle from introduction to maturity.
Source: Badiru (1992), own calculations.

Model implementation of technical change

When turning the attention to model implementation the historical context is interesting. The climate economics literature began to think about technical change early on, but it did not become a central theme for a long time. Since economist's IAMs originated from growth models, a natural point of integration was to switch over to endogenous growth theory. In an early version of such a model technical change results from carbon tax induced R&D (Goulder and Schneider 1999). General equilibrium theory proved particularly cumbersome with respect to this problem because of computational constraints and unstable solutions (Crassous et al. 2006). Quite astonishingly, Grubb et al. (2021b) writes that even though most models had incorporated some form of endogenous technical change by the time of the Fifth IPCC Assessment Report (2014), simulations were mostly run without it to save on compute time (see section 4.4 for some discussion). The computational complexity inherent in what amounts to introducing several differential equations into perfect foresight models led to rather ad-hoc approaches to the problem. Technical change was either incorporated through a research mechanism akin to endogenous growth theory or was applied to the overall energy efficiency of the economy

(Baccianti and Löschel 2014). These difficulties certainly contributed to the rise of partial equilibrium models ESM, focusing only on the most important economic sector in terms of carbon emissions.

Assuming that technology progresses has unequivocal consequences for mitigation policy: It lowers the cost of doing so. Less clear is the question of how far this carries and what the consequences would be in terms of policy. As the model base for the work of the IPCC grew in complexity and size, the question of technical change and how to deal with it got its own concerted effort with the Innovation Modeling Comparison Project (summarized in Crassous et al. 2006). This spawned a number of hybrid model approaches such as WITCH and REMIND that tried to make use of these findings and merge the top-down and bottom-up approaches of climate modeling (Farmer et al. 2015). With respect to climate policy, technical progress is the focal point of some key developments in the academic discussion. That early mitigation efforts, even though expensive, would eventually pay off, was raised by the Stern Review, for example (Weitzman 2007). It is one reason why decarbonization was re-framed as a systemic transformation, not a problem of a single market externality.

Table 3 summarizes how technical change is treated in the model sample. The most commonly used IAM, REMIND features learning-by-doing with cost floors. It also features learning for electric vehicles and electricity storage, making it the broadest implementation in the set. Learning does not extend to carbon capture or fossil fuels (Luderer et al. 2015). Treating technology as *exogenous* is apparently a deliberate decision for MESSAGE (Krey et al. 2020). Instead, a range of technology scenarios are prescribed – none of which match current evidence, as Way et al. (2022) have shown section 3. These constrain technology growth and improvements alongside the narratives of the IMPs for IPCC AR6 (2021). In IMAGE, learning parameters are so-called scenario "drivers". This means that they are standard inputs for the design of scenarios to model different trajectories (Stehfest et al. 2014). Scenarios may hence easily deviate from model assumptions. POLES also differentiates learning curves per technology and uses Wright's Law, but Keramidas et al. (2017) only refer to calibrations based on literature from an EU research project (ADVANCE). Witajewski-Baltvilks et al. (2015) was a core study in this project (see discussion in section 4.2), but it is unclear how this is reflected in POLES. The WITCH model is a product of the concerted efforts in the 2010s to improve technology representation in IAM (Crassous et al. 2006). It features differentiated learning curves per technology and a variety of backstops, but all technologies have cost floors (DeJong specification). It also has a very detailed energy CCS representation, aligning with past scenario prescriptions. Again, the model does not align better or worse with historic observations. A completely different approach is taken in GCAM. Despite featuring a reasonably detailed structure of the energy and power systems, technical change is not implemented at the level of technologies. Instead, the macroeconomic MAC is shifted to the right, effectively making a given level of abatement cheaper. It is unclear how this translates into detailed sectoral pathways, since this is the only way technological progress is documented in Bond-Lamberty et al. (2022). GEM-E3 bears the somewhat frustrating marks of the past, namely the need to account for macroeconomic consequences of energy price shocks, a macroeconomic hallmark borne out of the experience of the 1970s (Blanchard and Gali 2007). Learning-by-doing is implemented in the sense that the macroeconomy can reduce overall energy intensity (Capros et al. 2013). Reductions in *carbon* intensity on the other hand, which is ultimately what cheap zero emission technologies enable, is, per design, only possible by increasing energy efficiency.[24] Confusingly, GEM-E3 still features a detailed power sector representation, but much like in early approaches, it is entirely static in terms of economic trade-offs.

Table 3: Technical Change in Integrated Assessment Models

Model	Specification	Technologies
REMIND	learning-by-doing (DeJong)	Solar, Wind, BEV, electricity storage
	learning-by-waiting	Fossil fuels (transport, heat, power)
MESSAGE	learning-by-waiting	Exogenous technology pathways based on IMP narratives
IMAGE	learning-by-doing (Wright's Law)	Energy efficiency, fossil fuels and RES
POLES	learning-by-doing (Wright's Law)	Energy efficiency, fossil fuels and RES
WITCH	learning-by-doing (DeJong)	Solar, Wind, energy CCS[a]
	learning-by-searching (DeJong)	Energy efficiency, batteries (BEV)
	two-factor	Backstops (Nuclear, Oil, Industry)
GCAM	learning-by-waiting	Shift of the macroeconomic MAC
GEM-E3	two-factor	Energy efficiency

Formal representation of technical change in sampled models. In this context, RES refers to the four most important solar and wind technologies (PV and CSP, on- and offshore), which all models feature. Coverage of technologies is fairly limited. (a): Four different routes of fossil fuel power plants with CCS. Note that this does *not* include CDR technologies such as Bioenergy with Carbon Capture and Storage (BECCS) and Direct Air Capture (DAC). The former is often modeled separately for its sizeable implications on environmental systems, the latter is often a reference point for the least cost-efficient backstop. Sources: Model documentation, see table 2 for detailed references.

Overall, the possibility that progress in low-carbon energy supply technologies may be endogenous is well represented today, especially in the more commonly used models. Note that these are, without exception, deterministic representations. Whether a stochastic version of

[24] It is true that RES deployment can improve energy efficiency, but that primarily depends on the way primary energy is measured (see IPCC AR6 (2021, Annex I) and Brockway et al. (2021) for a discussion).

technical change would be more adequate for forecasting is another matter (a version of Moore's Law with error propagation is demonstrated in the appendix). This would account for the fact that technological development is additive in nature and builds on past successes and failures. But it adds computational complexity and using calibrated ranges (see for example Samadi (2018)) seems to be a useful simplification. The overall coverage of technologies is somewhat limited and power coupling and storage technologies, such as PtX and batteries are often not modeled. It is crucial to understand that neglecting progress in fossil energy technologies, which most models do, is not necessarily a helpful choice, as it could be a major source for rebound effects. Given adequate learning parameters and cost floors, models should be able to anticipate further progress, at least for the set of technologies represented here. The next question is whether these parameters are decently calibrated or not.

4.2 Calibration of Technical Parameters

If implementation of endogenous technical change is not the issue it could be a matter of calibration – in other words an empirical question. The various learning specifications discussed in section 4.1 can all be estimated with relative ease by using log-log models.[25] There are also some variations that have been proposed, such as industry scale (Goddard 1982) a combined Moore's and Wright's Law (Nordhaus 2014), or multivariate approaches (Nemet 2006). The literature doing these estimates is very extensive and spans almost a century. For obvious reasons, energy technology has taken center stage in recent years and Wright's Law or learning-by-searching (equation (5) and equation (9)) are popular models for the sake of their simplicity (Grubb et al. 2021a). Some conceptual theoretical frameworks have been proposed over the years to build a theoretical explanation for technology-specific learning by doing. Learning was understood in a more literal sense in early firm-level studies, so they linked it to labor conditions and worker churn (Yelle 1979). Sahal (1979) was the first to recognize that learning happens mostly in *complex* production processes as an aggregate result of finite and discrete changes in parts and routines. Goddard (1982) emphasized classic economies of scale of production plants. More recently, economies of scale are understood to play out in very different dimensions, so that the technology *itself* is either scalable or not (Farmer and Lafond 2016). Malhotra and Schmidt (2020) recently synthesized these ideas into a two-dimensional classification of (i) degree of

[25] In general, the term learning curve is used very loosely in the literature. I use it in the following as meaning any *endogenous* representation of technical change, not something akin to learning-by-waiting (equation (3)).

complexity and (ii) degree of customization. They find that a combination of high complexity and low needs for customization is most conducive to learning-by-doing, and propose that knowledge can be *embedded* in the production capital under these conditions. Overall though, the literature of decades over decades repeats that the subject is elusive by nature and conceptual frameworks can only give rough guidance.

Unsurprisingly, this situation has led to a lot of reluctance and mistrust in what little applied theory there is. There are simply severe conceptual issues with learning rate estimations that have arguably driven modelers to calibrate conservatively.[26] I will discuss the most important ones here, as they do intersect with the way models are constructed. I discuss these underlying empirical issues of learning rate estimation first. After that, section 4.2.1 and 4.2.2 discuss the actual calibrations. Since learning-by-searching implementation has, so far, not received consistent criticism it is left out here.

Omitted variables and lack of explanatory value

Equations (3) and (5) are basically univariate equations. This is obviously not ideal from a statistical standpoint: They leave no room for different hypotheses to be tested and have very limited value when interpreting results. Nemet (2006) criticized single-factor estimates for PV on the grounds that there are much more sophisticated and detailed models to be applied to readily available data. They propose controlling for commodity prices, module efficiency and plant scale, among others, to give a better appreciation of the exact nature of technical progress. This issue has been raised before: Sahal (1979) notes "worrisome ambiguities" as to "what is being learned and by whom". They show that the log-constant process implied in Wright's Law can be the aggregate result of several discrete and linear steps of progress in intermediaries.

In the context of IAM calibration this critique is somewhat misplaced, but seems to have driven some reluctance in accepting endogenous specifications (Grubb et al. 2021a). Since IAM are essentially forecasting tools, what matters is not causal representation but the accuracy of the relation that is being represented. In the case of technical change this means that it does not matter *what* drives progress as long as the predicted cost are accurate under a given scenario. The learning curve has a reduced form character in IAMs (Wiesenthal and Dowling 2012).

Witajewski-Baltvilks et al. (2015) expand this point further: The relation of factors influencing cost that are accounted for in models, and those that are not, have to be constant over time

[26] Strictly speaking, most of these critiques apply to Moore's Law as well, and even to fixed cost assumptions. Maybe it is better to have a bad theory of technical change than having none at all – but the problem just was not regarded in that way.

and over the scenarios created with the models. For estimation, this means that each model needs its own estimates instead of relying on previous literature. For example, a model utilizing two-factor learning curves, such as WITCH, would have to use estimates from a two-factor econometric model. This may seem like an obvious point, but it is not at all clear that models are calibrated like that. Luderer et al. (2015), Stehfest et al. (2014), and Keramidas et al. (2017) do not specify sources and instead point to other literature in general. The other model documentations in the sample just specify values. Only for WITCH there seem to have been custom studies (Drouet et al. 2019). Simultaneously, the lack of statistical rigor in the literature on technical progress has been raised repeatedly and may have veered modelers to calibrate conservatively (Nemet 2006; Nordhaus 2014; Grafström and Poudineh 2021).

Issues of data quality and interpretation

In almost all learning studies market prices are assumed to be a meaningful proxy for technology-specific progress. This is often a point of criticism (e.g. Samadi (2018)), but the meaning for modeling is never explicated fully. Market prices are a good cost metric precisely so long as markups stay constant over time. In standard microeconomic theory one would expect constant markups only in near perfect competition markets. Especially in early stages of a technology's adoption cycle that is strong assumption. On the contrary, interactions of market power with experience gains have been suspected as one of the drivers of early steepening of learning curves (Yelle 1979; Goddard 1982). The result of that would be a steepening of the cost progression, much like the Stanford-B representation in figure 2. No model from the IAM sample uses such a specification. It could be argued that learning-by-searching assumptions are used instead to cover the early stages of technology development (Drouet et al. 2019).

The two most prominent RES technologies, wind and solar, are not exactly new. Even though hardly any past study overestimated *deployment*, learning rates seem to have stayed somewhat constant for large enough panel data, with large variations between studies (Grubb et al. 2021a; Samadi 2018). As a caveat, there has been an acceleration this decade, discussed below. In any case, constant rates surely are not a good approximation for nascent technologies, that will play a major role in decarbonization, such as electrolysis, batteries or biofuels (Grubb et al. 2021a; IEA 2022).[27] There are other issues attached to the question of data selection of course, but again: What matters for a modeling perspective is that the reduced form relation is an accurate

[27] Part of the challenge here is of course that the expectations attached to certain technologies can depend on their potential. However, the technical requirements of net zero emissions and mostly intermittent energy sources do prescribe at least some residual roles to a certain set of supply side technologies.

representation. The larger question about cost metrics is discussed in section 4.3.

Reverse causality and policy influence

Cumulative demand can induce price reductions, but of course we would normally expect the reverse causality as well: Lower prices induce demand. If exogenous factors influence prices, this can lead to a biased estimator for the real learning-by-doing coefficient, as shown for example by Nordhaus (2014). A short formal exposition is given in the appendix. Note that this bias does not influence a baseline case without policy-induced demand, as we would simply expect the observed trend to continue. For the same reason, system-dynamic approaches such as Way et al. (2022), which simply prescribe historical trends, are not affected by this. The Nordhaus (2014) model also finds reasons for more ambiguous biases, but in general the concern was an upward bias in the learning rate. This may generally be justified, but we have seen that decarbonization scenarios have rather indicated the opposite conclusion (section 3). There are in fact well-founded reasons to dismiss overly cautious assumptions.

The first problem is that estimating a time trend and a learning rate simultaneously, which would alleviate reverse causality, leads to strong multi-collinearity. This is due to the fact that in most cases, production (and thereby experience, as measured in Wright's Law) increases exponentially over time, exactly mimicking the behavior of a linear time trend. Although not based on decarbonization technologies, Lafond et al. (2020) use a natural experiment approach, where time and experience are not correlated as strongly. As the US prepared national defenses in World War II, the federal government imposed what amounts to a price insensitive demand shock for military equipment on the economy, increasing defense spending by a factor of 30 in two years. As costs were secondary to victory, and spending was drastically reduced after the war, two demand shifts can be used to test for confounding exogenous factors. With this data set they find that an experience interpretation is robust to exogenous factors, but also that learning-by-doing varies strongly by the type of technology. Especially for scalable, large volume supply chains such as infantry equipment, exogenous factors were not as important.[28]

A second problem is that a core prediction of a combined model as in equation (12) is a *reduction* of the observed learning rate under deployment inducing policies. However, evidence shows that this may not always be the case. Wei et al. (2017) identify structural breaks in *learning*

[28] This is one of the many confirmations found in the literature that Goddard (1982) was not too far off the mark with the idea that economies of scale drive learning-by-doing. However, for similar reasons discussed here a scale specification is neither easily distinguishable from one over cumulative production, nor would it necessarily change calibration of IAMs.

curves to present evidence for the opposite: A policy induced steepening of the learning curve. With respect to solar energy, they find breaks in the US and Germany shortly after the introduction of tax credits and feed-in-tariffs, respectively.[29] Similar results are obtained across different technologies (lighting, fuel cells) and different types of regulatory changes (tax incentives, efficiency standards). There are very plausible reasons why policy could increase learning: If firms view policy measures as a signal for long-term support, they would ramp up R&D efforts. This would register in Wright's Law as a steepening. Another possibility is an induced scaling-up of production, that anticipates the future growth. Wei et al. (2017) do not test robustness or direction of causality however, so these findings are, to date, preliminary.

This positive interaction of policy and learning progress is generally not well studied with the only other recent example being Van Buskirk et al. (2014). This is somewhat surprising, since the idea that policy drives technical change is a central pillar in frameworks of green industrial policy (Rodrik 2014). Even more so, the study of learning-by-doing is in fact historically closely linked to industrial policy. Lafond et al. (2020) give a good overview of this earlier body of work, most of which focused on military technology. Some interesting questions that occupied these studies were for example whether capital accumulation or worker experience were of greater importance, whether R&D increased with demand and over contract designs, and how inter-plant spillover influenced cost improvements. Economists and political science scholars have recently begun to discuss a more planning centered approach to the energy transition (Krahé 2022; Grubert and Hastings-Simon 2022). Indeed, the focus of the learning literature on military technology is not a coincidence. More specifically, a planned approach could well be conducive to technical progress in key areas, and the historical context this literature emerged from suggests exactly that. We have to leave this as an open question here, but likewise it seems important to recognize that cost and technology dynamics were a much more prioritized issue in the post-war era of military industrial policy. This is for the simple reason that the heart of the matter, identifying characteristics *inherent* to technologies, is not left to the market and private actors alone, but of national interest. It is again, for the transition to a decarbonized economy. Robust empirical affirmation and theoretical exploration of such policy induced cost drivers could help in the context of effective mitigation policy design. More so, in the context of scenario modeling, learning rates endogenous to policy would represent somewhat of a paradigm

[29] Breaks are identified by application of the Akaike Information Criterion, which optimizes model errors over the degrees of freedom introduced. The results are 'event agnostic' breaks, exactly the opposite of what a fixed effects model, for example, would identify.

shift.

4.2.1 Calibration of Learning Rates

With all that in mind a look into the empirical literature and model calibrations is sufficiently contextualized. Even though the literature on technology diffusion goes back multiple decades, there is no evidence for adequate anticipation of RES trends, particularly for PV. Mirroring what has been shown in section 3, even single-technology empirical studies seem to have been consistently underestimating technical progress. This leads to the fairly puzzling through line that this pessimistic bias has been discussed extensively for decades now, without being thoroughly remedied. Most contributions raise reservations against a naive learning curve approach and that is of course warranted.[30] Refined approaches to the problem have recently emerged that may not give increased precision, but at least increase confidence in the applied methodology.

Older empirical studies on PV handily demonstrate the persistent difficulties in anticipating the successes of RES, so it is best to start off there. Note, that this is just to illustrate the issue – there are many other examples of early pessimistic studies, as Creutzig et al. (2017) review. Much prior to concerted efforts to improve technology representation in energy and climate models Grübler et al. (1999) point to the shortcomings of the largely static modeling stock in both the top-down and bottom-up traditions.[31] They observe learning rates in capacity cost for photovoltaics (20%), wind (20%) and gas turbines (10%) and show that cost reductions in PV have been very persistent since the 1970s. A "radical" scenario based on the MESSAGE model produces 100 GW of installed photovoltaics capacity in 2100. This is almost an order of magnitude off of the 843 Gigawatt (GW) actually installed in 2021 (IRENA 2021). Inflation adjusted, their projected module prices for 2100 are close to the global average total *installed* cost of today (830 versus 857USD/Kilowatt (kW), own calculations). Subsequent studies continuously validated the persistent improvements observed before, but remained conservative themselves. Without invalidating the disruptive progress underway, Nemet (2006) uses a multi-factor approach to gauge drivers of improvements. One scenario, where 30 Terawatt (TW) are installed in 2050 – entirely plausible given current market growth rates – reaches an inflation adjusted module price of 870USD/kW. The same can be said of a contribution by Wene (2008): Only

[30] With "naive" I refer here to a simple application of the univariate Wright's law.

[31] This is at a time when DICE-inspired macro models and sub-system dynamic models still largely existed next to each other and were not integrated to the degree they are today.

a hypothetical radical innovation scenario can just about match current prices. It is difficult to assess ad-hoc why this underestimation is so consistent, although a common theme is the worry about intermittency. The learning-rate literature frequently stresses the short-term character of the approach, something that does not align well with the needs of decarbonization scenarios (Malhotra and Schmidt 2020; Samadi 2018). However, it cannot be ruled out entirely that estimations are biased upward, too (see table 4).

Now, how do model calibrations compare to empirical results? A thorough meta-analysis of learning rate studies cannot be conducted here, much less over multiple technologies. Instead, I use two recent meta-studies from (Samadi 2018; Malhotra and Schmidt 2020) as a reference. Clear statistical support is found for the link between experience and cost for RES but also for fossil fuel technologies (the latter mostly single-digit percentages). Various multivariate estimations are common, especially controlling for R&D and commodity prices. None of the controls seem to systematically alter the results, but Samadi (2018) do not conduct a meta-regression.[32] Commodity price and market power controls have been frequent in recent literature, and improved statistical fit. A time control such as proposed by Nordhaus (2014) is also regularly done, but confirm the suspicion by Lafond et al. (2020) that neither fit nor central estimator are meaningfully improved in doing so. The most commonly used variables are installed capacity (dependent) and capacity cost (independent), presumably because of data availability. Again, studies using different metrics do not seem to produce meaningfully higher (or lower) learning rates or better fits. The exception here is wind power, where the highest results (learning rates of up to 32%) are reported for generation cost. Malhotra and Schmidt (2020) also collect and synthesize results, roughly in line with (Samadi 2018).

In general, model calibrations seem to align with meta-study recommendations for learning rates. However, it may be that conservative model calibrations are just a matter of inheritance from the learning literature itself. Again taking Samadi (2018) as a recent example, they summarize "plausible" results. But what is deemed so is left unexplained, and it is not at all clear if this involves a selection of the sample based on some plausibility criteria. For example, the aforementioned results for wind are out of what Samadi (2018) deem the acceptable range. It could be that this reflects the skepticism stemming from the empirical difficulties explained above. Besides that, the sampled studies often span very disparate time samples, some going back to the

[32] To the author's knowledge there are no meta-regression studies in the recent literature. This is somewhat surprising, but may be due to the methodological difficulties of properly controlling for varying and overlapping panel data in use.

Table 4: Learning Rates in Integrated Assessment Models

		Wind		Solar		CCS	Storage	Coal	Gas
		Onshore	Offshore	PV	CSP				
	high	12%	10%	23%	12%	12%	21%	5%	15%
Meta Analyses[a]	central	5%	3%	20%	8%	–	–	0%	6%
	low	-3%	-5%	15%	3%	2%	12%	-5%	2%
REMIND[b]		12%	12%	20%	-	-	10%	-	-
IMAGE[b]		20%	20%	20%	20%	-	N/A	5%	5%
WITCH[b]		10%	13%	17%	10%	3-6%	N/A	-	5%
Wright's Law (2010-21)[c]		38 ±10%	22 ±9%	38 ±5%	19 ±22%	-	–	-	-

Learning rates for different technologies from recent meta-analyses with ranges, model calibrations as per documentation and a "naive" first-difference application of Wright's Law on the most recent available global average generation cost data with standard errors. For nuclear energy, studies find consistently negative learning rates, which is also missing from model calibrations. But since nuclear does not feature heavily in mitigation scenarios, I omit it here. N/A means that the corresponding model does not feature a dedicated storage technology – WITCH and REMIND instead require dispatchable generation capacity when intermittent resources achieve high market shares. Storage values are based on a mix of battery chemistries and a mechanical short duration storage, similar to how REMIND models it. Way et al. (2022) uses similar calibrations to the ones implied in the meta-studies, but since all studies need to define a "synthetic" technology, I do not conduct an estimate for storage. Since there is no comparable data, I also omit estimates for Coal, CCS and Gas.
Sources: (a) Samadi (2018), Malhotra and Schmidt (2020), and Junginger and Louwen (2020), (b) model documentations (see table 2), (c) own estimates based on IRENA (2021) and equation (10).

1970s. Using this literature values, as documentations claim, then assumes that learning rates are invariant over time.[33] In contrast, early learning studies often treat Wright's Law as a short-term estimation tool. Hence, another reference is needed, if only to adjust to more recent data. For that purpose I use the following first-difference, fixed-effects consistent model proposed by Way et al. (2022) and Lafond et al. (2018), but I drop the error propagation, since it does not alter the estimator (see appendix 6). Crucially, I use LCOE data, *not* installation or capacity cost, from IRENA (2021) for the time period of 2010 to 2021. This has its problems too, but it is in line with recommendations of meta studies (Samadi 2018).

$$\log C_t - \log C_{t-1} = -\beta\left(\log y_t - \log y_{t-1}\right) + u_t \tag{10}$$

The left-hand side of equation (10) is the first-difference log of cost C_t, and the right-hand side is the first-difference *installed* capacity y_t.[34] Table 4 compares these results to those from

[33] For models that use a DeJong specification, such as REMIND and WITCH, this is a particularly curious juxtaposition, since they assume non-constant learning rates.

[34] This estimator takes care of omitted variable bias, but it should be kept in mind that I estimate over *installed capacity* instead of cumulative historical capacity. This is not uncommon in the literature (Junginger and Louwen 2020), but due to capacity depreciation learning rates are strictly higher when doing it this way. How the models implement this is not clearly explained, but given that no documentations expressly cite their initial deployment values one can assume that they use installed capacity as well.

the two meta-studies (Samadi (2018) and Malhotra and Schmidt (2020)) and the calibrations employed in REMIND, IMAGE and WITCH. The large deviations produced by the model in equation (10) has at least two reasons. The first one is the short time series, which coincides with a steepening of cost progression in wind and solar. In the case of PV and CSP this likely mimics in part the structural breaks identified in Wei et al. (2017), again without a definitive causal relation to policy changes. The second one is that it estimates over generation cost. The samples of Samadi (2018) and Malhotra and Schmidt (2020) contain too few studies doing the same, so it is not possible to say definitively whether this matches literature values, but section 4.3 discusses a few reasons why we would expect different values. It must be emphasized that this is not inherently a "better" estimate, but is intended as a point of reference. A reasonable calibration is another matter, but it is obvious from this that the model calibrations do not reflect recent cost progressions, and that meta-studies obfuscate this. Grubb et al. (2021a) examine some studies based on their sample time frame and confirm this "speed-up" for wind, but yet again only in studies specifying over capacity cost.

4.2.2 Calibration of Cost Floors

As previously discussed, REMIND and WITCH employ a DeJong specification of learning-by-doing, meaning that there is a cost floor below which technology cannot improve further. Recall from section 4.1 that the learning rate then asymptotically approaches zero.[35] This means that floor costs matter instantly in a simulation, such that the calibrated learning rate is never actually maintained. As a point of reference, table 5 shows the cost floors employed in both models in comparison to current global average data from IRENA (2021). Since all sampled models do not explicitly model battery or electrolysis storage via hydrogen or derivatives, these are not included here.[36] Similarly, for the backstops in WITCH and the vehicle prices in REMIND it is difficult to find meaningful comparisons. Both models seem to be more optimistic for large-scale energy systems, which strongly contradicts the general notion that scalability of production is what drives most cost progression (Samadi 2018; Lafond et al. 2018; Malhotra and Schmidt 2020).

Unfortunately there is no straightforward way to compare the "effective" learning rate re-

[35] Both Luderer et al. (2015) and Drouet et al. (2019) explicitly state that their cost curves are not linear approximations but indeed follow asymptotic flattening.

[36] Instead, they work with grid flexibility requirements or construct a synthetic storage technology (Luderer et al. 2015).

Table 5: Floor Cost in Integrated Assessment Models

	Wind		Solar	
	Onshore	Offshore	PV	CSP
	Assumed cost floors (USD/kW)			
REMIND	900	900	500	1300
WITCH	500	900	400	1500
	Actual total installed cost (2021, USD/kW)			
IRENA	1325	2858	857	9091

Cost floors employed in REMIND and WITCH, compared to current global average IRENA (2021). Values are not inflation adjusted, since model documentation does not explicate how the simulation starting point influences fixed prices. IRENA data depicts global average total installation cost, which is close to what models use (overnight installation cost, see Luderer et al. (2015)). Sources: See table 2 and IRENA (2021).

sulting from a DeJong specification to the values presented in section 4.2.1. Recall from equation (8), that the cost floor depends on the calibration of the initial cost C_0 and the compressibility factor θ. The effective learning rate on the other hand, when defined analogously to equation (6), would depend on the compressibility factor θ and cumulative deployment y_Σ. The effective learning rate therefore depends on C_0, which is not documented. Needless to say that the closer the cost floor to real-world observations, the stronger the dampening effect, and so effective learning in both models for PV and onshore wind is likely dampened.

The original DeJong model also had a theoretical foundation, the idea being that some production inputs do not benefit from learning-by-doing (Yelle 1979). Judging from the data presented in section 4.3 there could be some reason to deviate from a simple Wright's law for RES systems. Given that specific cost and yield vectors improved at different speeds, cost floors could approximate that. However, this would also imply using different learning rate calibrations than the overall rates commonly estimated in the literature, such as in (Samadi 2018). But even approaching the problem from this perspective is fraught with errors. Some authors suggest using multiples of the raw material input prices as a cost floor, since commodity prices are often stable over the long term and will take up an increasing share of overall cost (Junginger and Louwen 2020, p. 136). However, this entirely neglects substitution of materials or an increase in material efficiency.

In essence, there is little rigorous empirical support behind the ad-hoc assumptions on floor costs, and little in the way of sound methodology to rely on. Luderer et al. (2015) and Drouet et al. (2019) just specify values and these seem to be solely based on expert judgement. Given that this is also a common modeling practice outside the sample discussed here, it seems to be an issue that needs addressing on more thorough grounds (Way et al. 2022). As intuitive as they may be, to the author's knowledge there are no established methodologies to give accurate forecasts for cost floors. Identifying structural breaks in learning curves, such as done in Van Buskirk et al. (2014) and Wei et al. (2017) could be a way to identify flattening of technical progress. But biasing results against historical trends simply for the sake of intuitiveness seems a poor choice given the track record documented in section 3.

4.3 Cost Metrics and Scope of Change

The last section briefly touched upon another challenge when modeling technologies, which is how to actually measure cost. In economics, this problem is nothing unfamiliar: The links between market prices, costs and value are fuzzy and contested. Early in technology's deployment cycles this is of course exacerbated by price and market uncertainties. With respect to energy, more so to electric power, we might be lured by the textbook properties of the good: There are few more homogeneous products we could think of than charged electrons. However, energy technologies have undergone substantial changes in terms of their physical properties and the way they interact with one another. Ultimately there is no single, superior way to measure and compare technology performances, much less so over time (Way et al. 2022). This should be obvious: In whatever way we define a technology, it has multiple properties that decide in various ways over its utility in a wider context and system. But from a modeler's perspective this is a difficult spot to be in: There is no easy way of knowing beforehand how structural model decisions guide scenarios into all too rigid outcomes.

This is one reason why ESM modeling now commonly employ Agent-based Model (ABM) frameworks (for an early review, see (Weidlich and Veit 2008)). If reward structures and decision-making is more heterogeneous, the emerging opportunities for "useful" behavior must not rely on the social planner to understand them beforehand. Hoekstra et al. (2017) remarks that even performance metrics themselves are usually geared towards incumbent technologies and actors – one reason why "firm" generation and capacity provision might have taken a center stage in the modeling of IAMs (Drouet et al. 2019; Luderer et al. 2015; Krey et al. 2020). There is a long line of research into status quo biases and the role of expert opinions in reaffirming them. Agent based modeling approaches try to learn from that and may, in the future, be able to provide more comprehensive analysis than more rigidly structured IAMs (Farmer et al. 2015; Hoekstra et al. 2017). This is not the place to discuss these issues extensively, but is important to keep in mind two things: (i) The modeling approaches used in IAMs might be overly influenced by biases themselves and (ii) there may be proven modeling techniques that are able to solve issues surrounding innovation more comprehensively. Precisely because ABM structures are not as rigid in their decision-making processes they are deemed to be more successful in predicting network interactions in evolutionary, innovating systems (Farmer et al. 2015). For example, they have been successfully employed in predicting rapid BEV adoption in the past. In Shafiei et al. (2012) agents were allowed to have differing utility values about refueling ranges and be influenced by peers in their opinions towards vehicle types. Even in scenarios where gasoline prices do not increase[37] adoption rates follow exponential trends, closely matching empirical data in Iceland.

The back-and-forth between the IAM community and energy system modelers also has a long tradition (see section 4.4) and will probably continue to be fruitful in the future. Here, I can only take a more humble approach. I demonstrate how the measurement of performance and cost has influenced models, as well as the empirical studies they are based on, in the past. Due to a lack of in-depth studies for other technologies I only discuss these issues conceptually with the examples of wind and solar energy. The second part to this is how costs are actually modeled in comparison.

Installation and Generation Cost

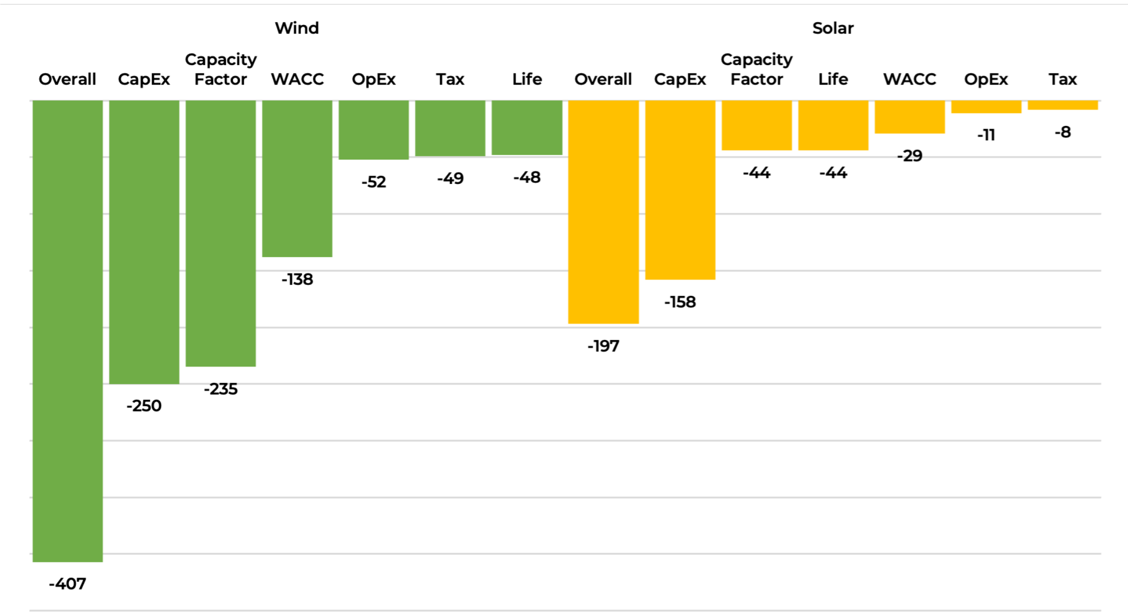
Recall that many learning rate estimates, as well as early meta studies, such as Creutzig et al. (2017) used installation cost and nominal capacities as performance metrics (sections 3 and 4.2). This is the reason why newer studies try to use generation cost in their calibrations (Way et al. 2022), particularly as it aligns better with the way models optimize scenarios. On the empirical side the most important reason for the use of installation cost is that these are readily available data directly facing the market. For example, almost all studies surveyed in Samadi (2018) on wind power use turbine prices as a cost metric, simply because it is the most accessible. Some do cross-checks with installation cost, which include construction and financing. But this misses major parts of the advancements in wind turbine design. More lightweight construction has allowed for longer blades and higher turbines, which enables the exploitation of both weaker and more stable wind resources (Bolinger et al. 2022). Performance and "viability" improvements like these are not visible in market prices, but instead improve the market *value* of the technology. In offshore wind installations, an increase in the installation costs has also been driven by the underlying economics, since moving farther offshore allowed access to better wind resources

[37] The price is of course *the only* driver in a typical IAMs

(IRENA 2021). The progress in wind turbines has therefore mostly resulted from an increase in the capacity factor, defined as the ratio of nominal capacity to actual average production.

For solar PV, improvements are less complicated, but learning studies mostly focus on module costs (Samadi 2018). This hides some significant improvements in module lifetime and an increase in the capacity factor of utility-scale installations, enabled by better control electronics (IRENA 2021). Victoria et al. (2021) expects capacity factor[38] improvements for PV in particular as the bulk of installed capacity slowly moves towards the equator (Victoria et al. 2021). Lifetime improvements could come from second use of degraded solar modules and is still completely understudied. For large-scale installations such as CSP, capacity factors are less of a reliable metric, as investments are typically very sizable. Raw project data such as reported by IRENA (2021) is not a good indicator of technical-inherent progress in that regard. (Bolinger et al. 2022; Ray 2021) use a harmonizing methodology to account for these factors, but the way this is done strongly depends on the technology in question and cannot be answered ad-hoc.

Figure 3: Decomposition of Cost Improvements in Renewables



Improvements (USD) of LCOE in wind turbines (1982-2020) and utility scale photovoltaics (2007-2020) in the US. For lack of data points, first years are averaged (wind: 4 years, solar: 3 years). Improvements in the capacity factor for wind have been driven by installation height, photovoltaics have improved in lifetime. Note that due to co-dependencies, factors are not completely additive. Source: Bolinger et al. (2022).

Figure 3 shows the harmonized results from Bolinger et al. (2022) for wind and solar power.

[38] the capacity factor relates energy output to nominal capacity, see below.

This kind of fine-grained analysis is difficult to conduct and typically only possible for very limited datasets (in this case for a short and recent time period in the US). One other major driver of cost improvements across both were the Weighted Average Cost of Capital (WACC). Globally, improvements of financing bottlenecks is a key priority in decarbonization policy particularly as renewable energy is relatively capital intensive (IEA 2023). As Bolinger et al. (2022) show, even in the US financing conditions for RES still improve, which is not always visible in ad-hoc assumptions about installation costs. The IEA for example has three different capital cost assumptions for the entire globe (IEA 2022).

For these reasons, newer empirical studies often use levelized cost metrics. A formal description can be found in the appendix – they provide a metric for the cost of electricity generation. This captures a wider range of dimensions technology evolution, but still must rely on generalizations for data gaps. To dispel the notion that this is a conclusive way of dealing with technology assessments, we can look at batteries and storage technologies. Again for reasons of data availability Way et al. (2022) treat batteries in the same ad-hoc manner as many learning rate studies and only apply learning to the price per storage capacity. By design, this *must* be an underestimate, since lifetimes will certainly improve in the future. More generally recent studies suggest severe limitations to levelized cost metrics when it comes to storage technologies, as they are highly sensitive to the system state. Parzen et al. (2022) propose to expand valuation based on market potential, which is designed to assess various kinds of interactions between grid design, placement of resources and storage times. In short, these are not issues that can conclusively be addressed by any modeling strategy.

Cost representation in Integrated Assessment Models

To understand how models feature technology characteristics, consider a stylized version of generation cost, with only one time period and without discounting. This is close to how REMIND models technologies, only that summation over the whole energy system happens first (Luderer et al. 2015). This may or may not alter the system outcome, but that is irrelevant for the question here.

$$C_E = \frac{\sum C}{\sum E} = \frac{I + O_t + V_t}{\delta E_P(t, \eta) T} \tag{11}$$

The generation cost are the sum total of installation cost I, the operation and maintenance cost O_t, typically zero for fuel-based production technologies, and variable cost (or fuel) V_t, always zero for RES. The total energy conversion output is given by the conversion efficiency η

times the annual primary energy input $E_P(t, \delta)$, times the total technology lifetime T. The energy input is a function of time for fuel-based technologies, and a function of the capacity factor δ, for intermittent sources such as RES. Table 6 summarizes how the models treat these different characteristics of energy supply technologies. In reality, none of these metrics are fixed for any technology, not even in global averages (IRENA 2021). However, the nuances play a more subtle role besides underestimating cost improvements. Except for IMAGE and REMIND, *all* models implicitly assume a slowing down of learning-by-doing, even when a Wright's Law specification is chosen. This is because Operational Expenditure (OpEx) are assumed to be constant and nonzero. As the technology progresses, a larger fraction of the cost is unchanging, analogous to the DeJong model (equation (7)). The former two models do not behave like that because fixed OpEx are either zero (IMAGE) or a fraction of the installation cost (REMIND). Improvements in *output* are ruled out in all models except GCAM and POLES. The former allows for exogenous progress and the latter assumes a progression path for wind turbines. It is not straightforward whether fixed performance assumptions have large effects. Changing technology characteristics can, in reality, make certain resources viable – the obvious case being wind turbines. But in a cost-driven model this could conceivably have the same effect as assuming the equivalent cost improvement – given that learning rates are calibrated to match *generation* cost. This would ultimately depend on the structure of the underlying supply curve and may only matter in edge cases.

Table 6: Technology Characteristics in Integrated Assessment Models

	CapEx		OpEx		performances		
	Installation[a]	Capital[b]	Fixed[c]	Variable[d]	Capacity Factor	Efficiency	Lifetime
	I	r	O_t	V_t	δ	η	T
REMIND	endogenous	constant	endogenous	market	constant	constant	"
MESSAGE	exogenous	"	constant	"	"	"	"
IMAGE	endogenous	"	0	"	"	"	"
POLES	"	"	constant	"	exogenous[e]	"	fixed
WITCH	"	"	"	"	constant	"	fixed
GCAM	exogenous	"	"	"	exogenous[f]	"	exogenous[f]

Technology characteristics in the sample IAMs. (a) Cost of physical capital (overnight), core learning-by-doing variable. (b) Regionally differentiated financing cost, pure cost of capital. (c) Operational and maintenance cost, only applies to RES. (d) Fuel cost, global fuel supply curves and carbon tax, or endogenous (biofuels). (e) Supply curve based on spatial resource quality, (e) Only applies to wind turbines. In POLES turbines grow over time, mirroring the development in figure 3. (f) GCAM allows for the prescription of gradual improvement in performance dimensions. Sources: Model documentations, see table 2.

Accounting for technical change on a per-variable basis is obviously no sensible choice given the sparse data and harmonization issues involved. But implying by design that technologies only change in terms of their installed cost has no theoretical or empirical basis at all. Overall, every model employs a very narrow scope of learning. The cost of finance is static in all models, despite showing learning effects in studies (figure 3), being expected to play a major role in emerging markets and developing countries (IEA 2022) *and* arguably representing a major policy lever. This poor representation of technology characteristics does not at all match recent evidence and urgently needs updating.

4.4 Interactions with Other Model Features

This section briefly discusses three issues that would arguably mostly influence deployment paths and not so much cost forecasts in scenarios. The general sense is that deployment forecasts have improved recently, but they could increasingly depend features that do not reflect recent evidence either IAMs (Way et al. 2022; Victoria et al. 2021; Bogdanov et al. 2021). The issues discussed here are very poorly described in model documentations, so the discussion here is kept brief and qualitative.

Non-constant marginal deployment cost
One component of what Grubb et al. (2021b) describe as dynamic realism is the idea of inertia, meaning that abatement cost increase with the amount abated in a given time period (see section 4). A plausible reason for this is that supply chains cannot ramp up instantaneously because they encounter macroeconomic bottlenecks. This is something models have incorporated, if maybe only for the purpose of smoothing out results: REMIND includes technology-level cost mark-ups in order to achieve "a more realistic phasing in and out of technologies" (Luderer et al. 2015, p. 21). This might seem like a reasonable abstraction, but the exact nature of these constraints is, once again, poorly documented compared to their potentially large role in simulation outcomes. It could, for example, be a step-wise function over the percentage increase in deployment.

But economic theory allows a few general remarks on such cost mark-ups. The main reason we would suspect costs to rise in this fashion are capacity constraints in production inputs: Physical capital, labor, intermediate inputs and raw materials. All these inputs are usually not modeled at all in IAMs, at least not in the sample discussed here.[39] In macroeconomic theory

[39] GEM-E3 features a partial representation of material flows and there are modular extensions for the other models as well (Capros et al. 2013). However, this usually serves the purpose of estimating the environmental footprint and resource needs of scenarios, not as a feature of the economic input-output structure.

capacity is obviously a very important, albeit ill-defined constraint. Schools of thought may disagree on whether a fast transition is subject to frictions, and where they originate from (for a recent EU-centered review, see Pollitt et al. 2017). There is good evidence that fears of inflationary drivers stemming from decarbonization are not warranted, and energy system studies like Way et al. (2022) are part of this evidence. But there is certainly room for reasonable disagreement.

Table 7: World Power Generation by Source

Source	Total Generation (PJ)			Annual growth rate	
	2000	2010	2021	2000-2010	2010-2021
Bioenergy	533	1166	2398	8%	7%
Oil	4219	3416	2736	-2%	-2%
Solar	4	112	3676	41%	37%
Wind	112	1246	6653	27%	16%
Nuclear	9025	9670	9857	1%	0%
Hydro	9443	12283	15214	3%	2%
Gas	9770	17071	22745	6%	3%
Coal	20588	30110	36400	4%	2%

World annual electricity production by source and respective implied growth rate (compound annual) show the radically different states the various production chains are in today. Source: Own calculations based on BP (2022).

On an industry level however, this is an entirely different matter: There is no reason to believe that any given supply chain cannot get used to sustained high growth rates. Firm level evidence suggests that certain sectors are better than others in absorbing shocks, notably those experiencing seasonal demand variations and those with complex supply chains (Cachon et al. 2007). But these effects are entirely unrelated to sectoral long-term productivity trends. Table 7 shows past growth trends in power generation by source. No economic theory would suggest that the overall *cost* trajectory of these technologies is affected by the speed of these trends. This can mean two things for the way IAMs model this "friction". If it is a function of absolute growth percentages for all technologies combined, there is an inherent bias against fast-growing power sources. Modelers may overcome that by calibrating on past per-technology growth rates. But of course growth trends can increase over the course of the energy transition, which is what we can likely expect from energy storage solutions (Battery and PtX). Alternatively, friction could be a function of *changes* in the growth rates. This would at last prevent *persistent* biases.

REMIND documentation makes direct mention of such constraints. Krey et al. (2020) vaguely mentions "technology diffusion constraints" for MESSAGE. It is unclear whether these are administered through cost or direct deployment limitations. The functional cost specifications of technologies in WITCH and IMAGE seem to work similarly to REMIND. For the other models there is no mention of such restrictions, but this could be due to poor documentation. Both Way et al. (2022) and Jaxa-Rozen and Trutnevyte (2021) mention these restrictions in a broader sample of IAMs. Whether these constraints show up in cost projections depends on how the model output is defined: If it is averaged over the existing capital, then scenarios would show an upward cost bias.

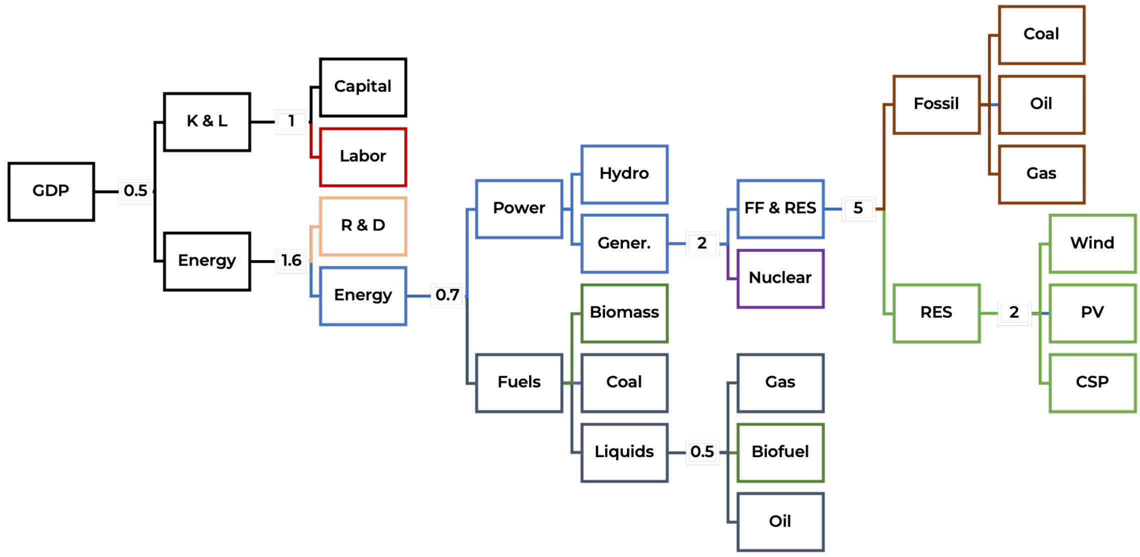
Nesting Structure, Substitution and System Constraints

The exact structure of models could matter quite a bit when it comes to technology adoption, especially in later stages of the energy transition (Sognnaes et al. 2021). Unfortunately this is also an area where model documentation is lacking. Only two models specifically demonstrate the underlying nesting structure, REMIND and WITCH (Luderer et al. 2015; Drouet et al. 2019). Instead of a typical cost-based Constant Elasticity of Substitution (CES) structure, other models, such as MESSAGE and IMAGE use flexibility coefficients for various technologies and then impose a target range for the overall energy supply (Stehfest et al. 2014; Krey et al. 2020). To entangle how these work in detail is not trivial at all, but some general remarks are still possible. For demonstration purposes, consider the structure of the WITCH model in figure 4 because it employs all types of constraints and is well documented.

A sensibility analysis by Carrara and Marangoni (2017) even mention that results in WITCH are highly sensitive to the nesting structure and values for the substitution elasticity. The nesting structure determines whether certain interactions are "physically" possible or not. For example, demand in WITCH is discrete and perfectly inflexible: There are no sectoral linkages downstream of the electricity supply. This is different in REMIND, where transportation is directly coupled to the power sector (Luderer et al. 2015). As to the elasticity, presumably a value of 2 is considered a common choice for energy technologies. But disagreements arise what to choose for fossil versus renewable sources. A sensitivity analysis for a baseline case finds generation shares of wind and solar between 2% and 30% for an elasticity of 2 and 10 respectively, in 2050 (Carrara and Marangoni 2017). These are very significant differences and should not be underestimated, particularly as they interact with learning dynamics.

WITCH also imposes two constraints, a flexibility and a capacity constraint. The former ensures a sufficiently stable supply in lieu of detailed modeling of temporal matching of power sources and demand. The latter is basically an energy security provision that only affects installed capacity, not generation. All three factors combined affect substantially the power output, system costs and installed capacity (Carrara and Marangoni 2017). Just as with learning rates, calibration estimates have to comply to the model structure. This is assuming that the model structure itself does not induce biases and instabilities, a phenomenon Thompson and Smith (2019) calls the *hawk moth effect* in the context of coupled climate-economy models.

Figure 4: Nesting Structure of the WITCH Model



Compacted structure, excluding backstops. Numbers refer to upstream elasticity of substitution – It is assumed to be infinite where not otherwise specified. All technologies are comprised of capital with vintage. Backstops are usually perfectly substitutable competitors to their counterparts (e.g. Coal versus Coal+CCS), only subject to capital vintage. In the REMIND model, technologies provision useful energy linearly and are not subject to a nested structure. They can be thought of as Leontief inputs instead, but the upstream technology is subject to substitution (e.g. electric and combustion cars). Source: Own (compacted) illustration based on Drouet et al. (2019).

More recent energy system studies make very different assumptions in the model structure to account for progress in storage technologies and demand flexibility (Bogdanov et al. 2021). This seems to mostly affect deployment results in the tail end, but it must be kept in mind that the system cost structure is *directly* co-determined by these assumptions – and so are the macroeconomic aggregates reported in IPCC scenarios.

Solution Algorithm, Optimization and Temporal Resolution

Another issue is the complex interaction of the way models find solutions and trace the temporal dynamics of scenarios. These links are impossible to entangle from one another, so I discuss them in tandem. On paper, they would also be a perfect question to explore in model sensitivity analyses, but to the author's knowledge, only REMIND allows for flexible time steps. For the others, temporal resolution is hard-wired to accommodate for the need of long-term projections

(Carrara and Marangoni 2017). A peculiar case is POLES, where the electricity sector is run on a two-hourly time step, while the others are running on a yearly basis. Default scenario increments are usually quite high between five and ten years (Krey et al. 2020; Luderer et al. 2015; Drouet et al. 2019; Stehfest et al. 2014).

The harsh trade-offs between structural realism, temporal realism, computation times and solution stability is quite well known in the modeling community, a phenomenon sometimes referred to as the *curse of dimensionality* (Maliar and Maliar 2014). In essence, the precision of solutions can drastically degrade when reducing time steps and solutions have to be approximated somehow. This is crucial for elements of the system where compounding occurs, which is the case for learning-by-doing dynamics. Cai et al. (2012) demonstrated with DICE that the ad-hoc approximation of IAMs, that treat the economy as a discrete time system produce large inaccuracies. For large-scale IAMs in use today, efficient numerical methods have to be applied. REMIND and MESSAGE documentation make mention of these problems and how they are addressed in principle, but it is not at all clear that accurate solutions are employed at the microeconomic scale. Given that discrete linearization can save heavily on computation time, even compared to efficient numerical methods (Maliar and Maliar 2014), this could be a trade-off modelers are willing to take on smaller scales. A bias against nonlinear relations and thus against fast-learning technologies would be the result. Grubb et al. (2021b) notes that in the last assessment cycle, many simulations were run without learning, exactly to save on computation time, so the modeling community has readily made these trade-offs in the past. ESM studies typically produce lower cost in model runs, as the trade-off is usually made in favor of temporal resolution, and they do not feature demand sector linkages at all (Way et al. 2022).

ESMs in general have had a back-and-forth relationship with IAMs. Structural assumptions in the latter about how to handle intermittent sources and how to link sectors with one another were based on more fine-grained energy system studies (Luderer et al. 2015; Carrara and Marangoni 2017). Above all, this concerns how grid integration of intermittent sources is approximated, since ESM can more readily analyze typical load structures and how to provision for them. But this has turned around in more recent studies that take a closer look at how sectors can be linked with one another since demand curtailment and load services can be provided by other sectors like industry and transport. In general, the conclusions about temporal resolution and the role of intermittent producers weaken significantly with sector linkages (Bogdanov et al. 2021; Shirizadeh and Quirion 2022). This does not mean that IAMs need a more complex structure and increased temporal resolution to yield better results. These are not issues that can be solved completely – rather they speak to the systemic nature of decarbonizing the world economy (Hoekstra et al. 2017). But structural calibrations need to be kept up to date, as their importance for policy grows.

4.5 Beyond Energy Supply: Demand Side Technologies

An underappreciated upside of the need for net zero is that it simplifies the choice of *how* to reduce anthropogenic emissions. In the climate science community there is little doubt that active CDR will be a requirement for stabilizing the climate (Allen et al. 2022; Baum et al. 2023).[40] There is no doubt that unabated fossil fuel use has to be stopped entirely, and that *abated* fossil fuel use has to be minimized to insignificant proportions in order to conserve limited carbon capture capacities (IPCC AR6 2022, ch. 1).[41] This includes non-energy use of fossil fuels in buildings ($\approx$ 5 Gt CO_2E), and transport ($\approx$ 8 Gt CO_2E). It further implies that industrial process emissions ($\approx$ 9 Gt CO_2E) have to stop entirely or be abated by CCS. This dramatically reduces the degrees of freedom for a decarbonized economy, which also applies to technology choices. The IEA track progress in key technological areas spanning across sectors and, by and large, all of them feature to some extent in any IPCC scenarios (IEA 2023; IIASA 2022)

It is no coincidence that ESM studies increasingly studied sectoral linkages, which is what this section will briefly discuss. What does technical change and the progress in RES technologies – which has been a focal point of this paper – mean for future scenarios and modeling work? From a systemic perspective the combination of economically viable non-dispatchable production and reliable storage splits up these two function of the energy system, thereby offering an increased number of ways to organize energy services (Way et al. 2022). A static representation of demand with grid integration mark-ups, which all the models employ, probably misses many of the nuances. This arguably has significant consequences for scenario-level aggregates.

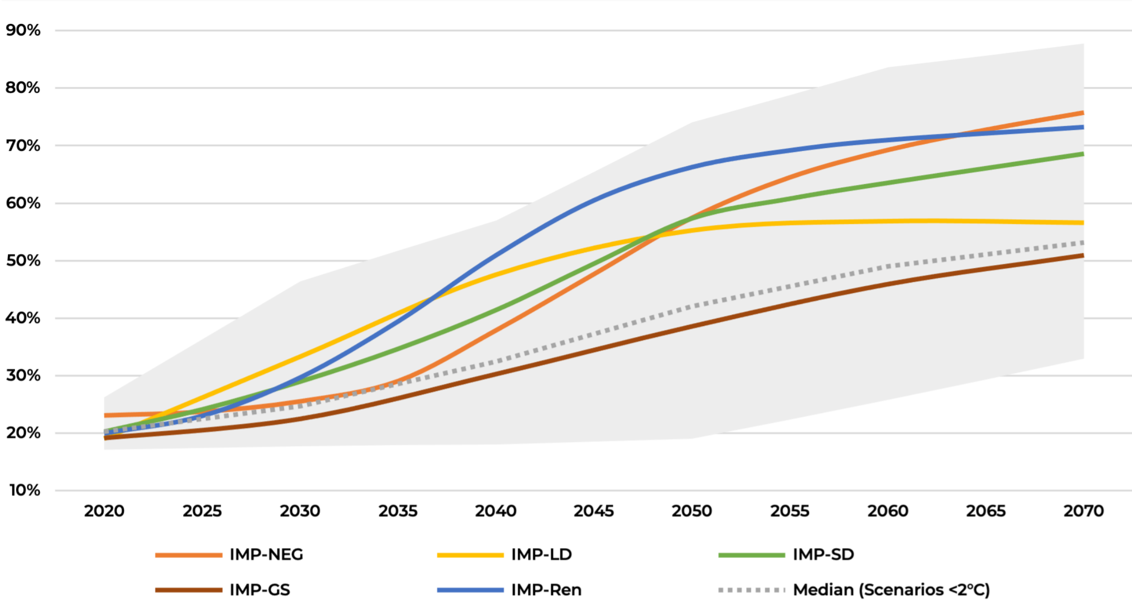
Figure 5 shows the electricity share in final energy demand projected in the IPCC AR6 (2021) IMPs and the scenario range.[42] As a rough point of comparison, a recent cost-efficient scenario by Bogdanov et al. (2021) sees electrification *in heating alone* at about 75% in 2050, higher

[40] The exact amount that will be needed is technologically statistically, and socioeconomically uncertain, but there is currently no Paris compliant IPCC scenario without Gigaton scale CDR (IIASA 2022).

[41] An edge case for fossil fuel use is the chemical industry, where some of it is used as a feedstock for production. This has varying implications for end-use emissions: Fertilizers are responsible for the majority of nitrous oxide emissions, while the worst environmental impacts of polycarbonates production are arguably not the residual emissions from production.

[42] A full narrative description fo IMPs can be found in IPCC AR6 (2021, ch. 1)

Figure 5: Electrification in IPCC Scenarios



Percent share of electricity in final energy demand for the Paris compliant Illustrative Mitigation Pathways. IMP-Neg: Focus on negative emissions. IMP-LD: Focus on demand reduction and efficiency. IMP-SD: Focus on alignment with sustainable development goals. IMP-GS: Focus on near-term alignment with Nationally Determined Contributions (NDCs). IMP-REN: Focus on rapid deployment of renewable energy. Grey Area: Maximum and minimum among Paris-compliant Scenarios. Source: IIASA (2022), own calculations.

than any scenario predicts for the whole power sector in IPCC AR6 (2021). Final energy is downstream of many conversion processes in industry, heating and transport, and so gives a rough picture of how much energy use is covered directly by electric power. A cheap power system, such as modeled in Way et al. (2022), is not anticipated in any of these and so would likely exceed these projections in a Paris-compliant pathway. This would have two ramifications: (i) It further narrows the band of critical technologies that play a major role in decarbonization, and (ii) it would induce positive feedback between demand and supply side technologies, given that learning effects can be exploited on the demand side as well.

With very few exceptions, the IAMs in the sample only feature rudimentary energy demand sectors. REMIND has a detailed transport sector module, but lumps together stationary energy demand sectors (Luderer et al. 2015). WITCH models industrial, residential and commercial non-electric energy demand as static backstops to fossil fuel combustion (Drouet et al. 2019). MESSAGE models all these sectors in a stylized way, with the energy sector providing carriers (fuel) with or without emissions. Historically, these were perfectly valid simplifications given the leverage of the energy sector. But they currently hide too much systemic interactions made possible by cheap decarbonized electricity.

Electrification of Transport

In transportation, the primary drivers of electrification will certainly be the adoption of BEV. Many IAM scenarios instead still see biofuels covering major shares of transportation energy, with shares ranging from about 10-40% in 2050 (IIASA 2022). This is an area that needs revisiting: Today the direction of the car industry and the competitiveness of synthetic fuels (derived from electricity) seem to be rather certain, reduce environmental trade-offs inherent to biofuels and enable much needed reuse of captured CO_2 (IEA 2023). The assumptions made for these sectoral couplings in IAM will very likely prove untenable. For example, in WITCH, BEV are subject to an ad-hoc learning-by-searching implementation, while Way et al. (2022) demonstrate strong learning-by-doing dynamics for batteries. REMIND covers this conceptually, but also implements a frictional term for switching, such as discussed in section 4.4. It is impossible to say how updating assumptions here would influence macroeconomic, or even sectoral outcomes. But again: given the strong feedback dynamics of electrified transportation, there could be sizable impacts. The sector would also continue to be a major driver of growth and rebound, matching the challenging trajectories in the past (IPCC AR6 2021, ch. 10).

Electrification of Industry

Industrial production is widely regarded as a major bottleneck for decarbonization, primarily because of the technical uncertainties and long-term capital investments necessary (IEA 2022). The latter is a major driver of technological lock-in and the policy guardrails for transitioning the sector will have to be defined sooner rather than later (Seto et al. 2016). Heavy industry is also a major driver of total energy consumption, and so the effects of electrifying industrial processes would be sizable. The problem is that many processes, such as metal smelting, ceramics and cement production mostly require heat, where traditionally, combustion processes have a large advantage over electricity. But a recent EU-centered study shows that almost 99 percent of industrial energy demand could technically be switched to electricity (Madeddu et al. 2020). Some of the alternative processes provide the co-benefit of being more conducive to secondary material routes enabling a circular economy. The key is that even though electrification is a technological option, the economics behind it are not fully clear yet. Neither the scheme of complexity and customization developed by Malhotra and Schmidt (2020) nor the notion of mass scalability would expect learning effects in these industrial processes. However, depending on the future cost of power, industry would provide additional demand and induce learning in the energy system itself. The impact on macro aggregates is not at all straightforward: In general, power-to-heat processes are very efficient and so can drive down the energy intensity of the economy (IPCC AR6 2022, ch. 11). But especially in the case of industrial capital this is not the result of smooth processes like efficiency learning implementations discussed above would suggest. Rather, these decisions lock in physical capital for upwards of 25 years (IPCC AR6 2022, ch. 11). Economic models typically would use what is called capital vintage or a putty-clay structure for this purpose, but none of the IAMs use this level of detail for industrial production. A generic demand sector will probably mask the opportunities for power sector coupling and the resulting learning effects (Luderer et al. 2015; Drouet et al. 2019; Capros et al. 2013; Stehfest et al. 2014).

Electrification of Heating

The potential for electrifying heating, mainly via heat pumps, has been discussed for decades. Empirical studies are limited, but arguably sufficient to characterize the potential at least roughly in integrated assessment models. Weiss et al. (2010) identifies two early studies that find learning rates of over 30% for residential heat pumps. More recently, Junginger and Louwen (2020) takes a more varied data set into account and finds slightly lower but comparable results of around 20-25%. Studies note similar technology-specific conceptual problems I have discussed here. But heat pumps are undoubtedly a key technology for decarbonization in the future and increasingly attract policy attention around the world (OECD and IEA 2022). Given that residential applications fit well into the criteria by Malhotra and Schmidt (2020) and production is scalable, there is no reason to doubt the empirical data. For IAMs scenarios, electrified heating could have major macroeconomic implications. OECD and IEA (2022) only project a 2% increase in electricity demand from a doubling of current heat market shares. But heating is considered one of the applications that could provide flexible demand and so lower system integration cost of RES (Bogdanov et al. 2021). The thermal work is mostly coming from the environment in heat-pumping systems, and so adoption will enhance primary energy efficiency by definition. Just as with industrial processes, past model projections could be vindicated because of that – but they would be correct in the aggregates for the wrong reasons.

5 Conclusion

Large-scale IAMs were initially never designed to inform climate policy in a comprehensive and detailed manner. Rather, their purpose was to find plausible, aggregate economic forecasts given the constraints of the climate, without too much regard for the underlying structure of the energy system (Way et al. 2022). The latter merely served as a guardrail. But as it turns out, technology assumptions have rendered aggregate trend predictions wildly inaccurate, and the IPCC assessment cycle is far too slow to address this in a timely manner. This, no less, at a time when IAMs are increasingly relied upon to devise detailed sectoral pathways to a net zero economy.

Some have called for IAMs to be scrapped entirely (Pindyck 2017), others have started work on expanding the model base (Hoekstra et al. 2017; Farmer et al. 2015; Dafermos and Nikolaidi 2022). The former is wildly impractical, the latter will hopefully prove successful in advancing our collective understanding of the challenges of decarbonizing the world economy. In the meantime, we are stuck with a set of highly complex assessment tools no single human can claim full understanding of (Ives 2021). Their importance will likely only grow. Critical examination of model output will be an important part of the scientific work surrounding decarbonization scenarios and well established tools to do so exist today (Koomey et al. 2019; Sognnaes et al. 2021).

Modelers, on the other hand, must do their part. Better and more comprehensive model documentation is urgently needed. Major flaws in the representation of the energy-economy system, such as discussed here, have to be addressed transparently. Modelers have presumably adjusted cost floors multiple times in the past but without any comprehensive addressing of the theoretical flaws (Way et al. 2022). Given the prominent role these models play, timely adjustments are important. But they cannot be conducted in such a manner that they perpetuate structural biases that have been well documented for close to a decade.

This paper has identified several drivers of poor technology forecasts in IPCC mitigation scenarios. Calibrations do not reflect current evidence, scope and scale of technical change are poorly captured, inertia assumptions and cost floors rest on shaky empirical grounds and structural assumptions make it impossible to capture major features of future energy systems. The best way to address these shortcomings would be a scenario architecture that puts far more weight on short-term realism, respecting historical trends. This is somewhat at odds with how IAM are used in IPCC mitigation assessments, but combining these two prospects seem like the most fruitful avenue for improvement in the utility of these policy tools.

In terms of mitigation policy, the consequences of this mismatch have roughly been spelled out in numerous other studies. Energy economy rebounds will be stronger than expected, with the associated advantages and disadvantages, and RES will be a strong driver of growth in the coming decades. The potential for electrification will prove to be large in many demand sectors. But as with RES before this will not happen on its own, at least not initially. The corresponding macroeconomic and wider social assessment is also quite straightforward. Much greater focus needs to lie on transitional dynamics and less on a comparative static between net zero and the economy of today. Above all, cheap low-carbon energy gives us urgently needed tailwind to achieve a stabilized climate in the middle of this century.

6 Appendix

Levelized Cost and Cost of Capital

The LCOE is the most commonly used levelized cost metric, designed to measure the cost of provision of end-use electricity for a given technology. It essentially represents what investors would have to be paid in order to provide a given amount of energy (often denoted in USD per Watt hour).

$$\text{LCOE} = \frac{C_T}{E_T} = \frac{\sum_{t=1}^{T} \frac{I+O_t+V_t}{(1+r)^t}}{\sum_{t=1}^{T} \frac{\delta \times E_n}{(1+r)^t}}$$

It is the total cost C_T over the total energy produced E_T over the lifetime T. Costs are comprised of the sum of all investment expenditure I_t annual operational cost O_t and variable, or fuel cost V_t. Electricity production is usually estimated based on a capacity factor δ and the nominal yield E_n of a technology. Both terms are discounted by r, often a risk adjusted and debt-to-equity weighted discount rate, the WACC. In practice, investors and power companies – unsurprisingly – use a variety of different approaches to assess risk that are not always based on formal, mainstream capital theory (Hürlimann et al. 2020). In contrast, IAMs and ESMs discount rate assumptions are typically very ad-hoc, and merely differentiate two or three world regions. For example, IRENA (2021) uses a value of 5% for OECD countries and China and 7.5% for the rest of the world. Levelized cost metrics are very sensible to discounting assumptions (IEA 2022).

There are substantial difficulties in harmonizing these cost metrics, and the technological characteristics as well as the exact accounting of source data can drive a lot of deviation. A strength of the data from IRENA (2021) are detailed and harmonized surveys on OpEx. Ray (2021) publishes an annual investor-focused report with a detailed breakdown of internal and external capital cost. Bolinger et al. (2022) normalize by taking commodities market data and macroeconomic conditions into account. As discussed in section 4.3, capital cost are partly a policy variable, because appropriate policy tools can reduce market failure. The LCOE itself is not "the" optimal measurement of economic value, and cannot capture fully the progress inherent to a given technology.

Causality Bias in Learning Rate Estimates

In the context of climate models, Nordhaus (2014) raised causality issues surrounding the learning curve literature. Given that most estimates relied on univariate specifications of Wright's Law, it is reasonable to ask whether it could be confounded by exogenous factors. For ease of exposition it is useful to refer to the formal model constructed by Nordhaus (2014). A combination of exogenous and endogenous learning, which is plausible ad-hoc for most technologies, would look like the following.

$$C_t = C_0 e^{-\alpha t} y_{\Sigma}^{-\beta} \tag{12}$$

This is just a combination of equations (3) and (5). The cost rate of change at t can then be given by

$$\hat{c_t} = \alpha + \beta \hat{y_{\Sigma}}, \tag{13}$$

where $\hat{c}$ and $\hat{y_{\Sigma}}$ denote the growth rates, respectively. If, additionally, demand grows over time and from substitution due to reductions in cost, output growth would follow

$$\hat{y_{\Sigma}} = \epsilon \hat{c}_t + d_t, \tag{14}$$

with ϵ as the (constant) elasticity of demand and d_t as any exogenous demand increase, for example due to income growth. Combining both it can be shown that the resulting estimator for the pure learning coefficient, $\bar{\beta}$, would look like the following.

$$\bar{\beta} = \frac{\hat{c}}{\hat{y_{\Sigma}}} = \frac{\alpha + \beta d_t}{\epsilon \alpha + d_t} \tag{15}$$

In this case, if we would estimate a pure univariate learning curve, the resulting $\bar{\beta}$ would confound exogenous learning, demand growth, substitution and the "real" underlying learning rate β. If the above model is assumed, bias from exogenous learning can be either positive or negative, depending on the elasticity ϵ. The elasticity itself plays a very different role, however. In this specification, for $\epsilon < 1$ it would bias the estimator upward, while for $\epsilon > 1$ it would bias it downward. Nordhaus (2014) was mostly concerned with the underestimation of mitigation cost resulting from biased learning assumptions.

Stochastic formulation of technical change

Farmer and Lafond (2016) show how error propagation can be used for a stochastic formulation of technical progress. A similar method is used in Lafond et al. (2018), Lafond et al. (2020), and Way et al. (2022), so it is worth demonstrating. The method is analogous for any of the specifications discussed in section 4.2, for example Moore's Law in equation (3).

$$C_t = C_0 e^{\alpha t} \tag{16}$$

The deterministic variant of Moore's Law above can be estimated with the following model:

$$\log C_t = \log C_0 + \alpha t + n_t, \tag{17}$$

where n_t is a noise term. In this case, noise cannot accumulate, meaning that two draws $C_{t1,t2}$ for any $[t1, t2]$ are always independent. But this would mean that a surprise discovery in a technology path is always offset by preceding or subsequent surprises in the other direction. Another possibility would be that any such deviations are only temporary fluctuations, such that there are no real discoveries about the underlying technology – only the steady march of predictable progress. Both interpretations are based on very strong assumptions. Instead, a random walk model can be used:

$$\log C_t = \log C_{t-1} + \alpha + u_t, \tag{18}$$

in which u_t is also a non-accumulative noise term. In the regression, n_t and u_t behave exactly the same, but the estimator for C_t is only dependent on C_{t-1} respectively. The iterative formulation, rewritten from 18 as a function of y_0 then allows for the noise to propagate.

$$\log C_t = \log C_0 - \alpha t + \sum_{i=1}^{t} u_i \tag{19}$$

Summation of the error terms u_i allows for previous shocks to persist. This model behaves like a true forecast: the further out the estimates are from any observed time frame, the larger the cumulative error becomes. Statistically, for any arbitrarily small error in the estimator of equation 17, the estimand leaves a given confidence range after a finite time frame. This is not the case in the random walk model. Treating technology forecasting as a stochastic process obviously adds another layer in computational complexity to current IAM, but it also stands in stark contrast to current modeling philosophy in general.

References

Allen, M. R. et al. (2022). "Net Zero: Science, Origins, and Implications". In: *Annual Review of Environment and Resources* 47.1. DOI: 10.1146/annurev-environ-112320-105050.

Arrow, K. J. (1962). "The Economic Implications of Learning by Doing". In: *The Review of Economic Studies* 29.3. DOI: 10.2307/2295952.

Baccianti, C. and A. Löschel (2014). *The Role of Product and Process Innovation in CGE Models of Environmental Policy*. Working Paper 68. WWWforEurope. URL: hdl.handle.net/10419/125722 (visited on 12/10/2022).

Badiru, A. (1992). "Computational Survey of Univariate and Multivariate Learning Curve Models". In: *IEEE Transactions on Engineering Management* 39.2, pp. 176–188. DOI: 10.1109/17.141275.

Baum, C. M. et al. (2023). *The State of Carbon Dioxide Removal. A Global, Independent Scientific Assessment of Carbon Dioxide Removal.* Edition 1. URL: bit.ly/3j9RdOL (visited on 01/22/2023).

Blanchard, O. and J. Gali (2007). *The Macroeconomic Effects of Oil Shocks: Why Are the 2000s So Different from the 1970s?* Working Paper 13368. Cambridge, MA: National Bureau of Economic Research. DOI: 10.3386/w13368.

Bogdanov, D. et al. (2021). "Low-Cost Renewable Electricity as the Key Driver of the Global Energy Transition towards Sustainability". In: *Energy* 227, p. 120467. DOI: 10.1016/j.energy.2021.120467.

Bolinger, M. et al. (2022). "Levelized Cost-Based Learning Analysis of Utility-Scale Wind and Solar in the United States". In: *iScience* 25.6. DOI: 10.1016/j.isci.2022.104378.

Bond-Lamberty, B. et al. (2022). *GCAM v3.2 Documentation*. URL: jgcri.github.io/gcam-doc/v3.2/toc.html (visited on 01/07/2023).

BP (2022). "Statistical Review of World Energy 2022". In: URL: on.bp.com/3IUNRd0 (visited on 01/17/2023).

Brockway, P. E. et al. (2021). "Energy Efficiency and Economy-Wide Rebound Effects: A Review of the Evidence and Its Implications". In: *Renewable and Sustainable Energy Reviews* 141.110781. DOI: 10.1016/j.rser.2021.110781.

Burn-Murdoch, J. (2022). "Economics May Take Us to Net Zero All on Its Own". In: *Financial Times*. URL: on.ft.com/3fflbyf (visited on 10/14/2022).

Cachon, G. P. et al. (2007). "In Search of the Bullwhip Effect". In: *Manufacturing & Service Operations Management* 9.4, pp. 457–479. DOI: 10.1287/msom.1060.0149.

Cai, Y. et al. (2012). *Continuous-Time Methods for Integrated Assessment Models*. Working Paper 18365. National Bureau of Economic Research. URL: ssrn.com/abstract=2143552 (visited on 01/27/2023).

Capros, P. et al. (2013). *GEM-E3 Model Documentation.* LU: European Commission. ISBN: 978-92-79-31463-6.

Carrara, S. and G. Marangoni (2017). "Including System Integration of Variable Renewable Energies in a Constant Elasticity of Substitution Framework: The Case of the WITCH Model". In: *Energy Economics* 64, pp. 612–626. DOI: 10.1016/j.eneco.2016.08.017.

Crassous, R. et al. (2006). "Endogenous Technological Change and the Economics of Atmospheric Stabilisation". In: *The Energy Journal Special Issue*. JSTOR: i23297001.

Creutzig, F. et al. (2017). "The Underestimated Potential of Solar Energy to Mitigate Climate Change". In: *Nature Energy* 2.9. DOI: 10.1038/nenergy.2017.140.

Dafermos, Y. and M. Nikolaidi (2022). *The Dynamic Ecosystem-FINance-Economy (DEFINE) Model: Manual*. Model Documentation. URL: define-model.org (visited on 10/14/2022).

Dietz, S. and N. Stern (2015). "Endogenous Growth, Convexity of Damage and Climate Risk: How Nordhaus' Framework Supports Deep Cuts in Carbon Emissions". In: *The Economic Journal* 125.583, pp. 574–620. DOI: 10.1111/ecoj.12188.

Dosi, G. and R. R. Nelson (2010). "Technical Change and Industrial Dynamics as Evolutionary Processes". In: *Handbook of the Economics of Innovation*. Vol. 1. Elsevier, pp. 51–127. DOI: 10.1016/S0169-7218(10)01003-8.

Drouet, L. et al. (2019). *Documentation – The WITCH Model*. URL: witchmodel.org/documentation/ (visited on 01/07/2023).

Farmer, J. D. and F. Lafond (2016). "How Predictable Is Technological Progress?" In: *Research Policy* 45.3, pp. 647–665. DOI: 10.1016/j.respol.2015.11.001.

Farmer, J. D. et al. (2015). "A Third Wave in the Economics of Climate Change". In: *Environmental and Resource Economics* 62.2, pp. 329–357. DOI: 10.1007/s10640-015-9965-2.

Georgescu-Roegen, N. (1981). *The Entropy Law and the Economic Process*. 4th ed. Cambridge (Mass.) London: Harvard university press. ISBN: 978-0-674-25781-8.

Goddard, C. T. (1982). "Debunking the Learning Curve". In: *IEEE Transactions on Manufacturing Technology* 5.4.

Goulder, L. H. and S. H. Schneider (1999). "Induced Technological Change and the Attractiveness of CO2 Abatement Policies". In: *Resource and Energy Economics* 1.21, pp. 211–253. DOI: 10.1016/S0928-7655(99)00004-4.

Grafström, J. and R. Poudineh (2021). *A Critical Assessment of Learning Curves for Solar and Wind Power Technologies*. OIES Paper EL 43. Oxford Institute for Energy Studies. URL: bit.ly/3rWbQOX.

Grubb, M. et al. (2013). *Planetary Economics: Energy, Climate Change and the Three Domains of Sustainable Development*. New York: Routledge. ISBN: 978-0-415-51882-6.

Grubb, M. et al. (2021a). "Induced Innovation in Energy Technologies and Systems: A Review of Evidence and Potential Implications for CO2 Mitigation". In: *Environmental Research Letters* 16.4. DOI: 10.1088/1748-9326/abde07.

Grubb, M. et al. (2021b). "Modeling Myths: On DICE and Dynamic Realism in Integrated Assessment Models of Climate Change Mitigation". In: *WIREs Climate Change* 12.3. DOI: 10.1002/wcc.698.

Grubb, M. et al. (2022). "Interactions of Time and Technology as Critical Determinants of Optimal Climate Change Policy". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3752788.

Grubert, E. and S. Hastings-Simon (2022). "Designing the Mid-transition: A Review of Medium-term Challenges for Coordinated Decarbonization in the United States". In: *WIREs Climate Change* 13.3. DOI: 10.1002/wcc.768.

Grübler, A. et al. (1999). "Dynamics of Energy Technologies and Global Change". In: *Energy Policy* 1.27, pp. 247–280. DOI: 10.1016/S0301-4215(98)00067-6.

He, G. et al. (2020). "Rapid Cost Decrease of Renewables and Storage Accelerates the Decarbonization of China's Power System". In: *Nature Communications* 11.1. DOI: 10.1038/s41467-020-16184-x.

Hoekstra, A. et al. (2017). "Creating Agent-Based Energy Transition Management Models That Can Uncover Profitable Pathways to Climate Change Mitigation". In: *Complexity* 2017.1, pp. 1–23. DOI: 10.1155/2017/1967645.

Hogan, D. et al. (2020). "Cost Estimating Using a New Learning Curve Theory for Non-Constant Production Rates". In: *Forecasting* 2.4, pp. 429–451. DOI: 10.3390/forecast2040023.

Hürlimann, C. et al. (2020). "Valuation of Renewable Energy Investments: An Explanatory Mixed-Methods Study about Applied Approaches amongst Practitioners". In: *Global Business and Economics Review* 22.1/2, p. 71. DOI: 10.1504/GBER.2020.105033.

IAMC (2022). *Common Integrated Assessment Model Documentation*. URL: iamcdocumentation.eu (visited on 10/18/2022).

IEA (2021). *World Energy Model Documentation*. URL: iea.org/reports/world-energy-model (visited on 10/17/2022).

IEA (2022). *World Energy Outlook 2022*. International Energy Agency. URL: bit.ly/3z0JYOi.

IEA (2023). *Tracking Clean Energy Progress*. IEA. URL: bit.ly/3H0Q1VR (visited on 01/23/2023).

IIASA (2022). *AR6 Scenarios Database*. Ed. by Byers, Edward et al. Version 1.1. Zenodo. DOI: 10.5281/ZENODO.5886911.

IPCC AR6 (2021). "Climate Change 2022: Mitigation of Climate Change." In: *Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by J. Skea et al. Intergovernmental Panel on Climate Change. URL: ipcc.ch/report/ar6/wg3.

IPCC AR6 (2022). "Climate Change 2021: The Physical Science Basis." In: *Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte et al. Intergovernmental Panel on Climate Change. URL: ipcc.ch/report/ar6/wg1.

IPCC SR15 (2018). *Global Warming of 1.5 ºC. 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, Sustainable Development, and Efforts to Eradicate Poverty*. Ed. by V. Masson-Delmotte et al. Intergovernmental Panel on Climate Change. URL: ipcc.ch/sr15.

IRENA (2021). *Renewable Power Generation Costs in 2021*. International Renewable Energy Agency. URL: bit.ly/3S7tKcw.

Ives, M. (2021). “A New Perspective on Decarbonising the Global Energy System”. COP26 event. URL: youtu.be/avnNwhs7gOA (visited on 11/08/2022).

Jamasb, T. (2007). “Technical Change Theory and Learning Curves: Patterns of Progress in Electricity Generation Technologies”. In: *The Energy Journal* 28.3, pp. 51–71. JSTOR: 41323109.

Jaxa-Rozen, M. and E. Trutnevyte (2021). “Sources of Uncertainty in Long-Term Global Scenarios of Solar Photovoltaic Technology”. In: *Nature Climate Change* 11.3, pp. 266–273. DOI: 10.1038/s41558-021-00998-8.

Jevons, W. S. (1865). *The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal Mines*. 2nd ed. London. URL: bit.ly/3btU80e.

Junginger, M. and A. Louwen (2020). *Technological Learning in the Transition to a Low-Carbon Energy System*. Elsevier. DOI: 10.1016/C2018-0-04547-8.

Keramidas, K. et al. (2017). *POLES-JRC Model Documentation.* LU: Publications Office. URL: doi.org/10.2760/225347 (visited on 01/06/2023).

Kesicki, F. and P. Ekins (2012). “Marginal Abatement Cost Curves: A Call for Caution”. In: *Climate Policy* 12.2, pp. 219–236. DOI: 10.1080/14693062.2011.582347.

Klepper, S. and E. Graddy (1990). “The Evolution of New Industries and the Determinants of Market Structure”. In: *The RAND Journal of Economics* 21.1, p. 27. DOI: 10.2307/2555491.

Koh, H. and C. L. Magee (2008). “A Functional Approach for Studying Technological Progress: Extension to Energy Technology”. In: *Technological Forecasting and Social Change* 75.6, pp. 735–758. DOI: 10.1016/j.techfore.2007.05.007.

Koomey, J. et al. (2019). “Inside the Black Box: Understanding Key Drivers of Global Emission Scenarios”. In: *Environmental Modelling & Software* 111.1, pp. 268–281. DOI: 10.1016/j.envsoft.2018.08.019.

Kouvaritakis, N. et al. (2000). “Modelling Energy Technology Dynamics: Methodology for Adaptive Expectations Models with Learning by Doing and Learning by Searching”. In: *International Journal of Global Energy Issues* 14.4. DOI: 10.1504/IJGEI.2000.004384.

Krahé, M. (2022). “The Whole Field”. In: *Phenomenal World*. URL: phenomenalworld.org/analysis/climate-planning/ (visited on 07/15/2022).

Krey, V. et al. (2020). “MESSAGEix-GLOBIOM Documentation - 2020 Release”. In: DOI: 10.22022/IACC/03-2021.17115.

Krey, V. et al. (2019). “Looking under the Hood: A Comparison of Techno-Economic Assumptions across National and Global Integrated Assessment Models”. In: *Energy* 172, pp. 1254–1267. DOI: 10.1016/j.energy.2018.12.131.

Lafond, F. et al. (2020). *Can Stimulating Demand Drive Costs down? World War II as a Natural Experiment*. MPRA Paper 100823. URL: mpra.ub.uni-muenchen.de/100823/.

Lafond, F. et al. (2018). "How Well Do Experience Curves Predict Technological Progress? A Method for Making Distributional Forecasts". In: *Technological Forecasting and Social Change* 128.1, pp. 104–117. DOI: 10.1016/j.techfore.2017.11.001.

Lamboll, R. et al. (2022). *Assessing the Size and Uncertainty of Remaining Carbon Budgets*. preprint. In Review. DOI: 10.21203/rs.3.rs-1934427/v1.

Lange, S. et al. (2021). "The Jevons Paradox Unravelled: A Multi-Level Typology of Rebound Effects and Mechanisms". In: *Energy Research & Social Science* 74.101982. DOI: 10.1016/j.erss.2021.101982.

Le Quéré, C. et al. (2019). "Drivers of Declining CO2 Emissions in 18 Developed Economies". In: *Nature Climate Change* 9.3, pp. 213–217. DOI: 10.1038/s41558-019-0419-7.

Luderer, G. et al. (2015). "Description of the REMIND Model (Version 1.6)". In: *SSRN Journal*. DOI: 10.2139/ssrn.2697070.

Madeddu, S. et al. (2020). "The CO_2 Reduction Potential for the European Industry via Direct Electrification of Heat Supply (Power-to-Heat)". In: *Environmental Research Letters* 15.12, p. 124004. DOI: 10.1088/1748-9326/abbd02.

Malhotra, A. and T. S. Schmidt (2020). "Accelerating Low-Carbon Innovation". In: *Joule* 4.11, pp. 2259–2267. DOI: 10.1016/j.joule.2020.09.004.

Maliar, L. and S. Maliar (2014). "Numerical Methods for Large-Scale Dynamic Economic Models". In: *Handbook of Computational Economics*. Vol. 3. Elsevier, pp. 325–477. ISBN: 978-0-444-52980-0.

Matthews, H. D. and K. Caldeira (2008). "Stabilizing Climate Requires Near-Zero Emissions". In: *Geophysical Research Letters* 35.4. DOI: 10.1029/2007GL032388.

Moore, G. E. (1965). "Cramming More Components onto Integrated Circuits [Reprint]". In: *IEEE Solid-State Circuits Society Newsletter* 11.3, pp. 33–35. DOI: 10.1109/N-SSC.2006.4785860.

Nagy, B. et al. (2013). "Statistical Basis for Predicting Technological Progress". In: *PLoS ONE* 8.2. Ed. by L. A. Nunes Amaral. DOI: 10.1371/journal.pone.0052669.

Nemet, G. F. (2006). "Beyond the Learning Curve: Factors Influencing Cost Reductions in Photovoltaics". In: *Energy Policy* 34.17, pp. 3218–3232. DOI: 10.1016/j.enpol.2005.06.020.

Nordhaus, W. (1992). *The 'Dice' Model: Background and Structure of a Dynamic Integrated Climate-Economy Model of the Economics of Global Warming*. Discussion Paper 1009. Yale: Cowles Foundation for Research in Economics. URL: bit.ly/3Rc7M96 (visited on 12/13/2022).

Nordhaus, W. D. (2013). *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World*. New Haven: Yale University Press. ISBN: 978-0-300-18977-3.

Nordhaus, W. D. (2014). "The Perils of the Learning Model for Modeling Endogenous Technological Change". In: *The Energy Journal* 35.1. DOI: 10.5547/01956574.35.1.1.

Nordhaus, W. D. et al. (1973). "The Allocation of Energy Resources". In: *Brookings Papers on Economic Activity* 1973.3. DOI: 10.2307/2534202.

OECD and IEA (2022). *The Future of Heat Pumps*. OECD. DOI: 10.1787/2bd71107-en.

Oswald, Y. et al. (2021). "Global Redistribution of Income and Household Energy Footprints: A Computational Thought Experiment". In: *Global Sustainability* 4. DOI: 10.1017/sus.2021.1.

Parzen, M. et al. (2022). "Beyond Cost Reduction: Improving the Value of Energy Storage in Electricity Systems". In: *Carbon Neutrality* 1.1. DOI: 10.1007/s43979-022-00027-3.

Pindyck, R. S. (2017). "The Use and Misuse of Models for Climate Policy". In: *Review of Environmental Economics and Policy* 11.1, pp. 100–114. DOI: 10.1093/reep/rew012.

Pisani-Ferry, J. (2021). *Climate Policy Is Macroeconomic Policy, and the Implications Will Be Significant*. Policy Brief 21-20. Peterson Institute for International Economics. URL: bit.ly/3uH9oOd.

Pollitt, H. (2019). *The Contribution of Post-Keynesian Economics to Climate Policy and Meeting Global Decarbonisation Targets*. URL: bit.ly/3sgIXOh (visited on 04/20/2022).

Pollitt, H. et al. (2017). *Technical Analysis on Capacity Constraints and Macroeconomic Performance*. Technical Study. European Commission. URL: bit.ly/3axPIpc.

Ray, D. (2021). *Lazard's Levelized Cost of Energy Analysis—Version 15.0*. Lazard Ltd. URL: bit.ly/3XX0kkI (visited on 12/05/2022).

Rodrik, D. (2014). "Green Industrial Policy". In: *Oxford Review of Economic Policy* 30.3, pp. 469–491. DOI: 10.1093/oxrep/gru025.

Romer, P. (1989). *Endogenous Technological Change*. w3210. Cambridge, MA: National Bureau of Economic Research. DOI: 10.3386/w3210.

Rosenbaum, E. (2019). "Rebound Effects and Green Growth – An Examination of Their Relationship in a Parsimonious Equilibrium Input-Output-Framework". In: *Journal of Cleaner Production* 225.1, pp. 121–132. DOI: 10.1016/j.jclepro.2019.03.296.

Rosslowe, C. and E. Cremona (2022). *New Generation. Building a Clean European Electricity System by 2035*. Ember Climate. URL: bit.ly/3YVDAm3 (visited on 12/28/2022).

Sahal, D. (1979). "A Theory of Progress Functions". In: *A I I E Transactions* 11.1, pp. 23–29. DOI: 10.1080/05695557908974396.

Samadi, S. (2018). "The Experience Curve Theory and Its Application in the Field of Electricity Generation Technologies – A Literature Review". In: *Renewable and Sustainable Energy Reviews* 82, pp. 2346–2364. DOI: 10.1016/j.rser.2017.08.077.

Semieniuk, G. et al. (2022). "Stranded Fossil-Fuel Assets Translate to Major Losses for Investors in Advanced Economies". In: *Nature Climate Change* 12.6, pp. 532–538. DOI: 10.1038/s41558-022-01356-y.

Seto, K. C. et al. (2016). "Carbon Lock-In: Types, Causes, and Policy Implications". In: *Annual Review of Environment and Resources* 41.1, pp. 425–452. DOI: 10.1146/annurev-environ-110615-085934.

Shafiei, E. et al. (2012). "An Agent-Based Modeling Approach to Predict the Evolution of Market Share of Electric Vehicles: A Case Study from Iceland". In: *Technological Forecasting and Social Change* 79.9, pp. 1638–1653. DOI: 10.1016/j.techfore.2012.05.011.

Shirizadeh, B. and P. Quirion (2022). "Do Multi-Sector Energy System Optimization Models Need Hourly Temporal Resolution? A Case Study with an Investment and Dispatch Model Applied to France". In: *Applied Energy* 305, p. 117951. DOI: 10.1016/j.apenergy.2021.117951.

Sognnaes, I. et al. (2021). "A Multi-Model Analysis of Long-Term Emissions and Warming Implications of Current Mitigation Efforts". In: *Nature Climate Change* 11.12, pp. 1055–1062. DOI: 10.1038/s41558-021-01206-3.

Stehfest, E. et al. (2014). *Integrated Assessment of Global Environmental Change with IMAGE 3.0 - Model Description and Policy Applications*. The Hague: PBL Netherlands Environmental Assessment Agency. ISBN: 78-94-91506-71-0.

Stiglitz, J. E. (1980). "A Neoclassical Analysis of the Economics of Natural Resources". In: *SSRN Journal*. URL: ssrn.com/abstract=250334 (visited on 05/21/2022).

Thompson, E. L. and L. A. Smith (2019). "Escape from Model-Land". In: *Economics* 13.1. DOI: 10.5018/economics-ejournal.ja.2019-40.

UNEP (2022). *The Closing Window. Climate Crisis Calls for Rapid Transformation of Societies.* Emissions Gap Report. United Nations Environment Programme. URL: unep.org/emissions-gap-report-2022.

Van Buskirk, R. D. et al. (2014). “A Retrospective Investigation of Energy Efficiency Standards: Policies May Have Accelerated Long Term Declines in Appliance Costs”. In: *Environmental Research Letters* 9.11. DOI: 10.1088/1748-9326/9/11/114010.

Victoria, M. et al. (2020). “Early Decarbonisation of the European Energy System Pays Off”. In: *Nature Communications* 11.1. DOI: 10.1038/s41467-020-20015-4.

Victoria, M. et al. (2021). “Solar Photovoltaics Is Ready to Power a Sustainable Future”. In: *Joule* 5.5, pp. 1041–1056. DOI: 10.1016/j.joule.2021.03.005.

Way, R. et al. (2022). “Empirically Grounded Technology Forecasts and the Energy Transition”. In: *Joule* 6.9, pp. 2057–2082. DOI: 10.1016/j.joule.2022.08.009.

Wei, M. et al. (2017). “Non-Constant Learning Rates in Retrospective Experience Curve Analyses and Their Correlation to Deployment Programs”. In: *Energy Policy* 107, pp. 356–369. DOI: 10.1016/j.enpol.2017.04.035.

Weidlich, A. and D. Veit (2008). “A Critical Survey of Agent-Based Wholesale Electricity Market Models”. In: *Energy Economics* 30.4, pp. 1728–1759. DOI: 10.1016/j.eneco.2008.01.003.

Weiss, M. et al. (2010). “A Review of Experience Curve Analyses for Energy Demand Technologies”. In: *Technological Forecasting and Social Change* 77.3, pp. 411–428. DOI: 10.1016/j.techfore.2009.10.009.

Weitzman, M. L. (2007). “A Review of The Stern Review on the Economics of Climate Change”. In: *Journal of Economic Literature* 45.3, pp. 703–724. DOI: 10.1257/jel.45.3.703.

Weitzman, M. L. (2009). “On Modeling and Interpreting the Economics of Catastrophic Climate Change”. In: *Review of Economics and Statistics* 91.1, pp. 1–19. DOI: 10.1162/rest.91.1.1.

Wene, C.-O. (2008). “Energy Technology Learning Through Deployment in Competitive Markets”. In: *The Engineering Economist* 53.4, pp. 340–364. DOI: 10.1080/00137910802482287.

Wiesenthal, T. and P. Dowling (2012). *Technology Learning Curves for Energy Policy Support*. JRC Scientific and Policy Reports. ISBN: 978-92-79-25676-9.

Wilson, C. (2012). “Up-Scaling, Formative Phases, and Learning in the Historical Diffusion of Energy Technologies”. In: *Energy Policy* 50, pp. 81–94. DOI: 10.1016/j.enpol.2012.04.077.

Witajewski-Baltvilks, J. et al. (2015). “Bending the Learning Curve”. In: *Energy Economics* 52.1, S86–S99. DOI: 10.1016/j.eneco.2015.09.007.

Wolf, M. (2022). “The Market Can Deliver the Green Transition — Just Not Fast Enough”. In: *Financial Times*. URL: on.ft.com/3UdGajV.

Wright, T. P. (1936). “Factors Affecting the Cost of Airplanes”. In: *Journal of the Aeronautical Sciences* 3.4, pp. 122–128. DOI: 10.2514/8.155.

Xiao, M. et al. (2021). “Plummeting Costs of Renewables - Are Energy Scenarios Lagging?” In: *Energy Strategy Reviews* 35. DOI: 10.1016/j.esr.2021.100636.

Yeh, S. and E. S. Rubin (2012). “A Review of Uncertainties in Technology Experience Curves”. In: *Energy Economics* 34.3, pp. 762–771. DOI: 10.1016/j.eneco.2011.11.006.

Yelle, L. E. (1979). “The Learning Curve: Historical Review and Comprehensive Survey”. In: *Decision Sciences* 10.2, pp. 302–328. DOI: 10.1111/j.1540-5915.1979.tb00026.x.