

WILLIAM USHER

THE VALUE OF LEARNING ABOUT CRITICAL
ENERGY SYSTEM UNCERTAINTIES

THE VALUE OF LEARNING ABOUT CRITICAL ENERGY
SYSTEM UNCERTAINTIES

WILLIAM USHER

20th July 2016 – v2.0.0

William Usher: *The Value of Learning about Critical Energy System Uncertainties*, An exploration of dynamic uncertainty in the energy system,
© 20th July 2016

DECLARATION

"I, William Usher, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis."

London, 20th July 2016

William Usher

ABSTRACT

In this thesis, a sensitivity analysis is used to systematically classify and rank parametric uncertainties in an energy system optimisation model of the United Kingdom, ETI-ESME. A subset of the most influential uncertainties are then evaluated in a model which investigates the process of resolving uncertainty over time — learning. The learning model identifies strategies and optimal pathways for staged investment in these critical uncertainties. By soft-linking the learning model to an energy system optimisation model, the strategies also take into account the system-wide trade-offs for investment across individual or portfolios of technologies.

A global sensitivity analysis method, the Method of Morris, was used to efficiently analyse the model over the full range and combination of input parameter values covering technology costs and efficiencies, resource costs, and technology/infrastructure build-rate and resource-constraints. The results of the global sensitivity analysis show that very few parameters are responsible for the majority of variation in the outputs from the model. These critical uncertainties can be separated into two groups according to their suitability for learning. Some of the important uncertainties identified, such as the price of fossil fuel resources available to the UK, are not amenable to learning and must be managed through risk-based approaches. The parameters which are amenable to learning, the availability of domestic biomass, and the rate at which carbon capture and storage technologies can be deployed, are then investigated using the learning model.

The learning model is formulated as a stochastic mixed-integer programme, and gives insights into the dynamic trade-offs between competing learning options within the context of the whole energy system. A UK case study shows that, if the resources are known to be available, total discounted net benefit of the availability of 150TWh/year of domestic biomass is £30bn, while the ability to build CCS plant at a rate of 2GW/year is worth up to £34bn. Together, the value increases non-linearly to a maximum of £59bn. This represents up to 17% of UK's discounted total energy system cost over the next four decades as quantified by the ETI-ESME model.

The learning model quantifies the cost threshold below which investment in an uncertain learning project is optimal. The threshold is a proxy for maximum no-regret investment over the aggregate total of research, commercialisation and deployment and could be of use to research funding agencies. The results show that when the likelihood of success of the project is 20%, one-stage learning projects of

£10bn or below are always undertaken. For the same likelihood of success, dividing a project into two-stages more than doubles the investment threshold to £22bn as it allows strategies in which investment switches away from a project if it fails. Dividing a project into multiple stages is particularly beneficial if most of the uncertainty is front-loaded, enabling switching to an alternative. The precise strategy to follow is a complex function of the cost, duration, net benefit and probability of success of each learning project, as well as the interactions between the project outcomes.

ACKNOWLEDGMENTS

A huge thanks to my wonderful wife, captain and inspiration, Dr. Bojana Bajzelj, who supported me, married me, and coaxed me to the finish line. Without her this would have been a less enjoyable and harder journey.

To the climbing crew, who kept me strong and healthy for past four years, the 8a challenge remains open: Ian Hogarth, Simon Atkins, Felix Ackermann, Stephan Knobloch, Sophie Parker and Bojana Bajzelj.

To the flatties, ex-flatties and extended-flatties thanks for the shoulders to lean on, the generous ears, the ready pints, the beds to sleep on, the nights-out and shared meals: Felix Ackermann and Mina Demiren, Somang Lee, Patrick Scholl, Monika Koziol and Basti Webb, Janina Ketterer and Stephan Knobloch, Ian Hogarth and Michelle You.

To my family, Sam, Charlie, Jan and Paul Usher. Sam thanks for many lunches and other mid-working-day distractions. Charlie for inspiration in music and culture, and my Mum and Dad for the continuing moral support and enthusiasm.

To all the staff at the UCL Institute of Making, particularly Rich, Zoe and Liz; your support and enthusiasm helped me through some dark days and founded a new passion for everything made with wood. On that note, Felix, the London Hackspace and PL[A]YWOOD hosted weekly entertainment throughout the latter years of my PhD and grew from strength to strength.

A huge thanks to all my colleagues and friends at the UCL Energy Institute who provided valuable advice and support, particularly: Steve Pye, Fabian Kesicki, Baltazar Solano-Rodriguez, Sophie Parker, Gabriel Anandarajah, Hannah Daly, Christophe McGlade, Will McDowall and Professors Paul Ekins, Tadj Oreszczyn and Bob Lowe. It's been an honour and a pleasure to work with and among such distinguished minds and genuinely lovely people. Its success is testament to my colleagues who work so hard, and the profound need for the research they undertake.

Finally, a very special thanks to Prof. Neil Strachan and Dr. Ilkka Keppo, my primary and secondary supervisors respectively. You pulled me through these four years and I thank you for getting me to the end. It was quite a journey. My examiners, Prof. Jim Skea and Dr. Laura Diaz-Anadon, provided insightful and learned comments, corrections and questions.

CONTENTS

Acronyms [xxii](#)

Models [xxii](#)

Nomenclature [xxiii](#)

1	INTRODUCTION	1
1.1	Essential Terminology and Concepts	2
1.2	The UK as a Case Study: Policies and Uncertainty	4
1.3	Uncertainties and the Energy System	5
1.4	Identifying Actions under Uncertainty	6
1.5	Combining a System Approach with Dynamic Uncertainty	7
1.6	Overview of this Study	8
1.6.1	Research Questions and Objectives	9
1.6.2	Academic Contribution	10
1.6.3	Thesis Outline	11
2	LITERATURE REVIEW	13
2.1	Treatment of Uncertainty in Energy System Modelling	13
2.1.1	Parametric and Structural Uncertainty in ESOMs	14
2.1.2	Formal Scenario Analysis	19
2.1.3	Monte Carlo Simulation	20
2.2	Sensitivity and Uncertainty Analysis	21
2.2.1	Sensitivity Analysis is not Uncertainty Analysis	22
2.2.2	Types of Sensitivity Analysis	23
2.2.3	Applications of Sensitivity Analysis to Energy Models	25
2.2.4	Use of sensitivity analysis in other fields	27
2.3	Modelling Approaches with Dynamic Uncertainty	28
2.3.1	Stochastic Programming	28
2.3.2	Stochastic Dynamic Programming	30
2.3.3	Stochastic Decision Trees	31
2.3.4	Selecting a Modelling Approach for the Learning Model	31
2.4	Learning: Reducing Uncertainty, Increasing Knowledge	32
2.4.1	Learning and R&D in the energy sector	34
2.4.2	Value of Information	36
2.5	Bringing all these aspects together	37
2.5.1	The Links Between Chapters	37
3	SENSITIVITY ANALYSIS: METHOD	39
3.1	The ETI-ESME Model	40
3.1.1	The Objective Function	41
3.1.2	Input Data	43
3.1.3	High-level Scenarios	44
3.1.4	Monte Carlo Feature	44

3.1.5	Key Decision Variables	45	
3.1.6	Spatial and Temporal Aggregation	45	
3.1.7	Elastic Demand Formulation	45	
3.2	Selection of an appropriate sensitivity analysis approach		46
3.2.1	Matching a sensitivity analysis approach to an energy system model	46	
3.3	Extending the Method of Morris to compute optimal input samples	51	
3.3.1	Optimal Trajectories	53	
3.3.2	SALib: A Python library for conducting sensitivity analysis	56	
3.3.3	Rationale for selecting distributions	56	
3.3.4	Grouping input parameters	58	
3.4	Integrating the sensitivity analysis, learning model and ETI-ESME	59	
3.4.1	Mechanism for treating the variation of inputs over time	59	
3.4.2	Mapping input uncertainties to agency	61	
3.4.3	Identifying meaningful output metrics	63	
3.4.4	Running a Morris analysis on an optimisation model	64	
3.5	Summary	66	
4	SENSITIVITY ANALYSIS: RESULTS	67	
4.1	Results from the Morris screening: Overview	67	
4.1.1	Interpreting the results	67	
4.1.2	Overview of the results	69	
4.2	Results from the Morris screening: Sectoral Investigation	75	
4.2.1	Emission Constraints	75	
4.2.2	Technology Costs	76	
4.2.3	Resource costs	77	
4.2.4	Build rate constraints	79	
4.2.5	Other parameters	79	
4.3	Including elastic demands	80	
4.4	Discussion	86	
4.4.1	Comparing Global Sensitivity Analysis to the Alternatives	88	
4.4.2	A Policy Perspective	90	
4.4.3	Learning about Uncertainty	91	
4.5	Summary	92	
5	LEARNING UNDER DYNAMIC UNCERTAINTY: METHODOLOGY	95	
5.1	Linking Research Projects and Dynamic Uncertainties		95
5.1.1	Modelling Research Projects	96	
5.1.2	The Structure of Research Projects	97	

5.1.3	Endogenous Uncertainty and Scenario Tree Structure	101
5.2	Learning Model	104
5.2.1	Outcome Generation	105
5.2.2	Implementing the Changes in Parameter Values in the ESME model	106
5.2.3	Incorporating the ESME revenue function into the R&D model	110
5.2.4	Introducing Non-Anticipativity Constraints into the Deterministic ESME Scenarios	111
5.3	Exploring the Learning Model: One Research Project	112
5.3.1	One Stage, Two Time-Periods	113
5.3.2	Two Stage, Two Time-Periods	115
5.4	Structuring Multi-Stage Research Projects	116
5.4.1	Mutually Exclusive Research Projects	117
5.4.2	Mutually Exclusive Projects - A Portfolio	119
5.4.3	Under Perfect Additionality (Independent Projects)	119
5.4.4	Imperfect Additionality (Partially-Independent Projects)	121
5.4.5	Extension to Probabilities <1	121
5.5	Summary	122
6	LEARNING UNDER DYNAMIC UNCERTAINTY: RESULTS	123
6.1	Case Study	124
6.1.1	Choice of Parameters	124
6.1.2	Defining the Case Studies	125
6.1.3	Revenue Function: Partial Revenues	128
6.1.4	Revenue Function: Big Pharma	129
6.1.5	Learning Model Outcomes	130
6.1.6	Investigating Strategies	130
6.2	One-stage Results	134
6.2.1	Individual Projects - One-stage	134
6.2.2	Both Research Projects - One-stage	136
6.3	Two-stage Results	140
6.3.1	Two stage, "big-pharma"	140
6.3.2	Two stage, partial-benefits	147
6.4	Energy System Model Results	152
6.4.1	Availability of Domestic Biomass - Two stage	154
6.4.2	Build rate of CCS Plant - two stages	159
6.4.3	Combining Biomass Availability with CCS Build Rates - One stage	159
7	DISCUSSION	169
7.1	Answering the Research Questions	170
7.1.1	What is the best method to identify and rank uncertainties?	170

7.1.2	What is the value of learning about critical energy system uncertainties?	171
7.1.3	What is the structure of dynamic energy-system uncertainties?	171
7.1.4	How do dynamic energy-system uncertainties relate to opportunities for learning?	172
7.2	Implications for Policy	172
7.2.1	Results from the Learning Model	172
7.2.2	Using ESOMs to Inform the Energy R&D Agenda	176
7.2.3	Policy implications of the Global Sensitivity Analysis	178
7.3	Implications for Research	178
7.3.1	Reflections on Alternative Methodologies	179
7.3.2	Model Insights from the Global Sensitivity Analysis	180
7.4	Limitations and Future Work	183
7.4.1	Sensitivity Analysis	183
7.4.2	Evaluation of Critical Uncertainties	185
7.4.3	The Learning Model	185
7.4.4	Extensions to the Learning Model	187
7.4.5	Energy Research Projects	190
7.4.6	Eliciting Uncertain Input Data for the Research Projects	195
A	THE ELEMENTARY EFFECTS METHOD	197
A.1	Sampling Method	197
A.2	Working with Groups	198
A.3	Computing Elementary Effects	198
A.4	Computing Metrics	198
B	THE STOCHASTIC MIXED INTEGER PROGRAMME	201
B.1	Decision variables	201
B.2	Non-anticipativity constraints	202
B.3	Other constraints	202
C	UNCERTAINTY FRAMEWORKS	205
D	INPUT DATA	209
D.1	Grouping of input parameters for sensitivity analysis	209
E	ADDITIONAL RESULTS: OPTIMAL TRAJECTORIES	215
F	ADDITIONAL RESULTS: SENSITIVITY ANALYSIS	217
G	ADDITIONAL RESULTS: LEARNING MODEL	221
G.1	Predicted values for the Revenue Function	221
G.2	Strategies from Two-Stage “Big-Pharma” Projects	221
G.3	Strategies from Two-Stage “Partial-Revenues” Projects	225
	Glossary	229
	BIBLIOGRAPHY	237

LIST OF FIGURES

- Figure 1 A diagram of a small portion of the ESME model. The solar resource (red) is consumed in a solar collector conversion technology (yellow) to generate a hot water resource (purple). The hot water is converted in Dwelling conversion technologies (yellow) to produce the hot water energy service (orange) to meet the energy service demands in dwellings of varying densities (grey) 41
- Figure 2 This figure shows the effect upon the randomly generated input sample to a model of the number of levels p as the number of trajectories N increases. 52
- Figure 3 This figure shows the effect upon the input sample to a model, without (top four rows) and with (bottom) the optimisation procedure. The optimisation avoids the worst incidences of sampling bias, such as for the second variable in the fourth row. 55
- Figure 4 Example simulation capital cost trajectory for different samples of ω 60
- Figure 5 Influence on total cost of 31 inputs groups of parameters 70
- Figure 6 Influence on carbon price in 2050 of 31 inputs groups of parameters 71
- Figure 7 Influence on electricity production of 31 inputs groups of parameters 73
- Figure 8 Influence on nuclear capacity of 31 inputs groups of parameters 74
- Figure 9 Influence on proportion of electricity capacity that is renewable in 2050 of 31 inputs groups of parameters 74
- Figure 10 Sensitivity of total cost (in the reference scenario without elastic demands and with no emissions constraint) 81
- Figure 11 Sensitivity of total cost (under the policy scenario with elastic demands) 82
- Figure 12 Sensitivity of change in consumer/producer surplus (under the policy scenario with elastic demands) 83

- Figure 13 Sensitivity of carbon price in 2050 (under the policy scenario with elastic demands) 85
- Figure 14 A schematic showing the linkages between the ETI-ESME model and the learning model 96
- Figure 15 A tree that represents a two-stage research project 98
- Figure 16 The value of flexibility increases with the number of stages T over a range of probabilities p of research project success. The cost of completing the research project $C = 10$. 100
- Figure 17 Scenario tree: Binary results, non-sequential projects 102
- Figure 18 Scenario tree: Binary results, sequential projects 102
- Figure 19 A decision tree that represents a four stage research project, with equivalent (continuous) cumulative probability distribution. The true probability distribution would be discrete. 103
- Figure 20 The decision tree associated with one research project split into two stages. If the first stage succeeds (1_s), the decision maker can continue to stage two, in which case the set of results is $(2_s, 2_f)$. 115
- Figure 21 Optimal strategy under mutually exclusive revenue function 118
- Figure 22 Optimal strategy under mutually exclusive revenue function with hedging 120
- Figure 23 Optimal strategy under perfect additionality 120
- Figure 24 Results of the multiple regression 129
- Figure 25 Some example strategies. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively. 132
- Figure 26 The normalised objective function of the learning model given the change in normalised cost of the biomass availability project with a 10% discount rate, and with an artificially extreme decrease in revenue - £30bn in 2030 and £15bn in 2040 135
- Figure 27 The normalised objective function of the learning model given the change in (undiscounted) cost of the biomass availability project with a 10% discount rate, while a CCS build rate project is fixed at 10% likelihood and a high cost of £20bn 136

- Figure 28 Four plots corresponding to probabilities of 20% or 80% for success of the research projects for biomass availability and CCS build rate 138
- Figure 29 In these plots, the cost of the second stage (not included in the value on the axes) for both projects was fixed to £10bn. The cost of the first stages for both projects was varied up to £15bn and is displayed on the x- and y-axes for CCS and biomass availability respectively. While the probability of the project was fixed at a 20% likelihood of succeeding, the balance of the probabilities over the first and second stages are altered across the four plots to demonstrate the effect upon the strategies selected. The strategies numbered here bear no relation to those numbered in Figure 28. 141
- Figure 30 Six of the twelve main strategies from the example two-stage big-pharma projects, continued in Figure 31. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively. 142
- Figure 31 Six of the twelve main strategies from the example two-stage “big-pharma” projects, continued from Figure 30. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively. 146
- Figure 32 The pattern of strategies under different combinations of uncertainty front-loading with partial-revenues. The numbers refer to the strategies pictured in Figures 33 and 34 148
- Figure 33 Six of the eleven main strategies from the example two-stage partial projects. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively. 149
- Figure 34 Five strategies from the 0.25/0.80 balance of probabilities with a two-stage project with interim revenues shown in Figure 32a. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively. 151
- Figure 35 The scenario tree for biomass availability (two-stages, three key-years). The tree for the one-stage research project with three key-years is a sub-tree of this one, removing those nodes marked with a *. 155

Figure 36	The change in installed capacity (GW) over time as a consequence of biomass availability	156
Figure 37	The change in resource use (TWh) over time as a consequence of biomass availability	157
Figure 38	The change in net CO ₂ emissions (MtCO ₂) over time as a consequence of biomass availability	158
Figure 39	The change in installed capacity (GW) over time as a consequence of build rate of CCS	160
Figure 40	The change in resource use (TWh) over time as a consequence of build rate of CCS	161
Figure 41	Annual net CO ₂ emissions (MtCO ₂) over time as a consequence of build rate of CCS	162
Figure 42	Annual change in capacity over time as a consequence of build rate of CCS and biomass availability	163
Figure 43	Annual change in resources over time as a consequence of build rate of CCS and biomass availability	164
Figure 44	Annual change in net emissions over time as a consequence of build rate of CCS and biomass availability	165
Figure 45	A schematic of energy system components according to novelty and granularity for research purposes	191
Figure 46	Methodological responses to different forms of uncertainty (adapted from Stirling, 2007, 2010)	206
Figure 47	Framework for handling uncertainties and their effects (adapted from McManus et al., 2006)	207
Figure 48	From a pool of trajectories, examples of the best case samples	215
Figure 49	From the same pool of trajectories as above, examples of the worst case samples	215
Figure 50	An initial screening analysis indicated that some parameter groups were influential.	217
Figure 51	A more detailed analysis of the top 18 groups of parameters (of 44 tested) to which the ETI-Model is most sensitive.	218
Figure 52	Results with carbon constraint set to 80% reduction by 2050	219
Figure 53	Strategies 1 to 6 from the two-stage big-pharma results in Section 6.3.1	221
Figure 54	Strategies 7 to 12 from the two-stage big-pharma results in Section 6.3.1	224
Figure 55	Strategies 13 to 15 from the two-stage big-pharma results in Section 6.3.1	225

Figure 56	Strategies 1 to 6 from the two-stage partial projects in Section 6.3.2. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.	226
Figure 57	Strategies 7 to 12 from the two-stage partial projects in Section 6.3.2. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.	227
Figure 58	Strategies 13 to 18 from the two-stage partial projects in Section 6.3.2. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.	228

LIST OF TABLES

Table 1	Energy System Modelling Studies 1979 to 2012	16
Table 1	Energy System Modelling Studies 1979 to 2012 continued	17
Table 1	Energy System Modelling Studies 1979 to 2012 continued	18
Table 2	A summary of the global sensitivity analysis techniques applied in energy related studies	26
Table 3	List of inputs to the ESME model	43
Table 4	Breakdown of the categories of inputs in the Energy System Modelling Environment (ESME) model, the number of parameters in each category, and the subset of parameters available for uncertainty quantification. The total available uncertain parameters for sensitivity analysis numbered around 200.	44
Table 5	Characteristics of sensitivity and uncertainty analysis methods (adapted from Flechsig et al., 2012; Saltelli et al., 2008a) given an upper bound on CPU time (3 days)	47
Table 6	Mapping uncertain inputs to agency. DD indicates whether the uncertainty can be modelled as decision dependent - using the learning model in Chapter 5.	62
Table 7	Model output indicators	64
Table 8	Ranking of technology cost input groups	76
Table 9	Ranking of resource cost input groups	78
Table 10	Ranking of max build rate input groups	79
Table 11	Ranking of efficiency input groups	80

Table 12	Typology of Uncertainties for Decision Dependent problems 97
Table 13	The ESME scenarios enumerate the outcomes across the key years 109
Table 14	The structure of the relationship between outcomes and scenarios, enforced by the constraint shown in Equation 45. The final two rows enumerate the results and key years associated with each of the six scenarios. 111
Table 15	Scenarios for R&D projects. Numbers in brackets are capital costs for the (sub-optimal) investment.. These scenarios correspond to those shown in Table 13. 114
Table 16	List of research projects 116
Table 17	The structure of the relationship between outcomes and scenarios, enforced by the constraint shown in Equation 45. Only outcomes 2,3,8,9 have a probability > 0 with the inputs shown in Table 16. F=Fail, P=Pass. 117
Table 18	The revenue is a function of the observed combination of outcomes (of which there are nine permutations) at the end of the time horizon. 118
Table 19	A more relaxed revenue function does not penalise the hedging approach and so allows combinations in which stage 1 of the non-successful project was completed successfully. 119
Table 20	Under perfect additionality, revenues are summed, so successfully completing both projects (scenario 9) results in a doubling of revenue over completing just one project successfully. 120
Table 21	Example revenues under imperfect additionality 121
Table 22	The cases and location of the results in this chapter 125
Table 23	Research levels for domestic biomass in the one-stage case 126
Table 24	Research levels for domestic biomass in the two-stage case 126
Table 25	Research levels in the one-stage case 126
Table 26	Research levels in the two-stage case 127

Table 27	The probabilities of the research project outcomes in the learning model. The probabilities for each column sum to one (values are rounded to two decimal places). These probabilities are ‘invented’ to demonstrate the learning model and bear no relation to the probability distributions contained in the ESME data set. 130
Table 28	The ten viable strategies for CCS Build Rate and Biomass Availability. The numbered key corresponds with that of Figure 28 . 139
Table 29	The 15 viable strategies for CCS Build Rate and Biomass Availability with two-stage big-pharma research project. The numbered key corresponds with that of Figure 29 . 144
Table 30	Sectoral emissions in the base year (2010) of the ETI-ESME model 153
Table 31	Storage Costs 209
Table 32	Resource inputs 210
Table 33	Constraints 210
Table 34	Heating Technology Costs 211
Table 35	Technology Costs and Efficiencies 212
Table 36	Transport inputs 213
Table 37	The modelled and predicted values for the revenue function 222
Table 38	The revenue function for the “big-pharma” result set. 223

ACRONYMS

- EMF** Energy Modelling Forum. 16, 37
- EPSRC** The Engineering and Physical Sciences Research Council is a funder of fundamental scientific and engineering research in the UK. 5
- ETI** Energy Technologies Institute. A public-private partnership consisting of Shell, BP, EDF, E.On, Caterpillar, Rolls Royce and the UK Government. ETI is responsible for conducting research into energy technologies. xiv, 5, 13, 22, 38, 41, 42, 46, 50, 66, 89, 93, 95, 98, 150, 167, 171, 173, 175
- ETSAP** Energy Technology Systems Analysis Program. xiv
- IAM** Integrated Assessment Model. xiii, xiv, 13, 16, 22, 26, 31, 33, 35, 174, 202, 203
- IEA** International Energy Agency. xiv
- IIASA** International Institute for Applied Systems Analysis. xiv
- IPCC** Intergovernmental Panel on Climate Change. xiii
- RCP** A set of scenarios of future high-level drivers (such as population growth and economic prosperity among others) of climate change used to inform and underpin the work of the [Intergovernmental Panel on Climate Change \(IPCC\)](#). 20
- RCUK** From the website <http://www.rcuk.ac.uk>: Research Councils UK (RCUK) is the strategic partnership of the UK's seven Research Councils. Each year the Research Councils invest around Â£3 billion in research covering the full spectrum of academic disciplines from the medical and biological sciences to astronomy, physics, chemistry and engineering, social sciences, economics, environmental sciences and the arts and humanities. 5, 6
- UKERC** The UK Energy Research Centre (UKERC) carries out world-class research into sustainable future energy systems (from <http://www.ukerc.ac.uk>). 5

MODELS

- DICE** The Dynamic Integrated Model of Climate and Economy (DICE model) is an [Integrated Assessment Model \(IAM\)](#).. 27, 31, 34
- PRICE** Implements a probabilistic version of the DICE model. 36

- ESME** The Energy System Modelling Environment was developed for [Energy Technologies Institute \(ETI\)](#) by Redpoint, an energy consultancy. 5, 9–11, 13, 22, 38, 41–44, 46–48, 50, 52, 53, 60, 63–66, 68, 69, 71, 74, 77, 81, 89, 90, 92–95, 97, 98, 106, 108, 110, 112–114, 117, 123, 125–127, 130, 131, 135, 137, 147, 150, 152, 154, 167–169, 171, 173–175, 177, 178, 180–186, 224
- GCAM** The Global Change Assessment Model is an integrated assessment model, which models the interactions between energy, agricultural, economic, land-use and technology systems. It has evolved from the MiniCAM model documented in a series of papers starting with (Edmonds et al., 1983). Full details of this open-source model can be found on the website https://wiki.umd.edu/gcam/index.php/Main_Page. 26, 27, 35
- MARKAL** *Market Allocation*. MARKAL is a generator of energy-system models, not an [Energy System Optimisation Model \(ESOM\)](#) in its own right. Developed under an implementing agreement of the [International Energy Agency \(IEA\)](#) called [ETSAP](#). xiv, 14–16, 27, 74
- MESSAGE** Developed by [International Institute for Applied Systems Analysis \(IIASA\)](#). 14, 15
- MIND** Model of Investment and Technological Development. 33
- TIMES** The Integrated [MARKAL](#) EFOM System. Developed by [ETSAP](#). Successor to [MARKAL](#). 64
- WITCH** World Induced Technical Change Hybrid model. WITCH is an [IAM](#) designed to evaluate the impacts of climate policies on global and regional economic systems and to provide information on the optimal responses of these economies to climate change, see <http://www.witchmodel.org> for more details. 27

NOMENCLATURE

Sets

- (i, j) The compound set of project [stages](#), page 105
- Ω_i set of possible result for project i, page 106
- $i \in I$ The set of projects, page 105
- $j \in J$ The set of [stages](#), page 105

$s \in S$ The set of outcomes, page 105

$z \in Z$ Set of scenarios in the ESME model denoted by the combination of years t chosen in the ESME model, number of technologies i and number of steps j . See Equation 40, page 110

$k \in K$ Set of resources, page 203

Subsets

\hat{z}_z^{my} Non-anticipativity scenario from which to copy values for scenario z , page 112

$d_{s,z}$ Legal scenarios z that can be selected for outcome s , page 110

$e_{z,t^{KY}}$ Legal scenarios z in key year t^{KY} , page 111

I_{is}^S Technologies i that succeed in scenario s , page 201

$i^{ss'}$ The differentiating technology for scenarios s and s' , page 202

$j^{ss'}$ The differentiating research stage for scenarios s and s' , page 202

oc_{si}^F subset of result combinations where the result is a failure in outcome s , page 106

oc_{si}^P subset of result combinations where the result is a success in outcome s , page 106

$oc_{si} \in OC$ set of result combinations in outcome s for project i , page 106

R_{sij} Identifies the scenarios s in which stage j is successful for project i (excludes the final stage if successful), page 201

t_z^{my} Non-anticipativity threshold year for scenario z , page 112

Decision Variables

Obj Objective function (total expected net revenue), page 201

$W_{sz} \in \{0, 1\}$ Binary variable which indicates whether ESME revenue at combination z is selected for outcome $s \quad \forall (s, z) \in d_{s,z}$, page 110

$X_{ijts} \in \{0, 1\}$ 1 if research and development (R&D) project (i, j) starts at the beginning of period t in outcome s , page 106

$Y_{ijts} \in [0, 1]$ 1 if R&D project (i, j) is complete at the beginning of period t in outcome s , page 106

$Z_{ijts} \in [0, 1]$ 1 if R&D project (i, j) can be started at period t in outcome s , page 106

Financial Variables

- Cost_{ij}^R Cost of **R&D stage** j for project i , page 201
- Cost_s^T total development cost in scenario s , page 201
- $d\text{Rev}_{it}$ discounted revenue at time t for project i , page 201
- h'_z Discounted revenues for each scenario z , page 109
- h_z Total energy system cost from **ESME** model for scenario z , page 109
- Rev_s^T total revenue from successful completion of **R&D** projects in scenario s , page 201

Parameters

- \hat{p}_{ij} Probability of **R&D** project **stage** i, j succeeding, page 106
- ρ_{ijk} Resource requirements, page 203
- τ_{ij} duration of **R&D stage** i, j , page 106
- b_{zit} Computed parameter which contains all the possible combinations of research project **stages** j for each project i and time-period t in scenario z , page 110
- l_i **Stride** of project i , page 106
- p_s probability of **outcome** s , page 106
- r Discount rate, page 201
- $\text{sdif}_{ss'}$ Scenarios differ by one, page 202
- ρ_k^{\max} Resource availability, page 203

INTRODUCTION

Energy policy is fundamentally concerned with managing [risks](#) and [uncertainties](#) to energy, economic and social systems. Energy policy in the UK is often framed in the form of the *energy trilemma*: the need to ensure that energy is sustainable, secure and affordable (Rhodes et al., 2014). For example, the move to a *sustainable* energy system is about managing the risk of global climate change. Energy *security* is concerned with ensuring that economic productivity (i.e. the continuous consumption of energy for productive ends) can be maintained in the face of global economic and geo-political uncertainty, and increasingly, the engineering challenges associated with integrating the increasing proportion of variable electricity sources into the National Grid. And *affordability* ensures that UK industries remain globally competitive and that social equity is maintained through access to reasonably priced energy. Affordability is a function of geo-political and wider economic issues, investment decisions and operation of the energy market.

There are a number of critical uncertainties which could undermine one or more of these three policy dimensions. Determining which uncertainties are important is not a trivial task, as almost everything is uncertain to some degree. Adopting an appropriate framework within which to model uncertainties could give some insights. One such framework is the use of [Energy System Optimisation Models \(ESOMs\)](#). The past decade has seen a surge in the use of numerical models applied to various isolated components of the energy system, as well as for analysing the interactions across the whole system. However, as explored in Chapter 2, these models have not yet been used to systematically identify critical uncertainties in the energy system.

Even if it were possible to pin-point critically important uncertainties, the challenge remains to determine what actions can be taken in the face of uncertainty. The normative approach would have decision makers resort to risk-management, with a balanced [strategy](#) determined through weighting the costs and benefits of different outcomes according to their likelihood of occurrence. However, such an approach treats all uncertainties in the same way: as static and unchanging. A volatile oil price is managed in the same way as the availability of a critical low-carbon technology. A local wind-speed forecast is shoe-horned into the same approach as the changing winds of global economic growth. In short, traditional risk-management techniques ignore many important aspects of uncertainty. And it is not

clear what sort of approach should be used to manage uncertainty across the energy system.

Under some circumstances, decisions can influence the nature of uncertainty, or its probability, or hasten its resolution. One key action, or intervention, policy makers can make is to invest in research and development. Research and development projects can be viewed as an example of managing [dynamic uncertainty](#). [Dynamic uncertainties](#) are when uncertainties change over time as we learn more and increase knowledge. Through funding a [research project](#), a decision maker can find answers to unresolved questions¹. Within the framework of this thesis, research and development is an example of an approach to manage or reduce uncertainties, a process otherwise referred to more generally as *learning*.

The rationale for approaching energy policy and energy system modelling from the perspective of [dynamic uncertainty](#), is to attempt to bridge the gap between model outputs and the needs of policy makers. If policy makers wish to make energy policies that are not only evidence-based, but robust, adaptive or flexible, then the evidence-base must reflect this. Investigating how investments in learning — the resolution of uncertainties — can influence future decisions could be extremely valuable to policy makers. Existing frameworks for combining knowledge into models of the energy system fall short of this goal.

The remainder of this chapter explores these issues in more depth. The next Section [1.1](#) presents some essential terminology. Sections [1.2](#) to [1.5](#) briefly outline the problem areas that are tackled in this study. Section [1.6](#) gives an overview of the thesis, including the research questions and objectives, the academic contribution of this study and an outline of the thesis.

1.1 ESSENTIAL TERMINOLOGY AND CONCEPTS

The term [uncertainty](#) appears almost 500 times in this thesis. In isolation, [uncertainty](#) is used in the general sense, encompassing all states of incomplete knowledge ([Knightian uncertainty](#), [risk](#), [ambiguity](#), [ignorance](#)) as set out in Stirling, (2007, 2010) and reproduced in Appendix C.

More often, [uncertainty](#) is prepended with a modifying adjective. [Aleatory uncertainty](#) and [epistemic uncertainty](#) refers to the differing natures of incomplete knowledge. In the former case, the uncertainty is something that is irreducible, while the latter is reducible.

¹ This requires careful phrasing of a research question. For example, a more specific question such as “What developments would be required for tidal stream to produce electricity at less than £80/MWh by 2030”, is more useful to answer than a vague question such as “Is tidal stream technology cost effective in the UK?”

The term **dynamic uncertainty** refers to **epistemic uncertainty** which changes over time as a result of investment, experimentation or waiting. A **dynamic uncertainty** is a target for **learning**. In contrast a static uncertainty is one that is fixed and cannot be learnt about or changed. The terms static and dynamic are used to describe a (potentially subjective) perspective on uncertainty. In contrast, the terms aleatory and epistemic describe fundamental natures of uncertainty.

For example, the wind speed outside my window is an aleatory uncertainty. In existing energy system models, technology cost is treated as a static uncertainty, although the uncertainty is in fact epistemic and could be reduced through learning. A new model could treat technology costs as dynamic uncertainties, acknowledging the link between investments in R&D, the uncertain outcomes of R&D and the corresponding reduction in technology costs.

The modelling community use the technical terms **endogenous uncertainty** and **exogenous uncertainty** to distinguish between the specific mechanisms which affect the learning process. In the former, modelled decisions affect either the timing of the learning process or the success of the learning process². In the latter case, the timing of the resolution of uncertainty is not affected by the model, and occurs irrespective of what happens in the model. The more explicit term **decision-dependent uncertainty (DDU)** means the same as **endogenous uncertainty** but for optimisation models - where a **decision variable** affects the outcome of an uncertainty.

These three categories of pair-wise definitions are related as follows. The terms aleatory/epistemic refer to the fundamental (and objective) nature of an uncertainty. An uncertainty can be composed of both aleatory and epistemic components. In contrast, dynamic/static refer to a subjective perspective on the same uncertainties. However, an aleatory uncertainty is never dynamic, while an epistemic uncertainty is never static. Endogenous/exogenous refer to how an uncertainty is modelled. Static (aleatory) uncertainties may only be modelled exogenously. Dynamic (epistemic) uncertainties should be modelled endogenously, but are typically only modelled exogenously.

Later in the thesis, I use three terms to distinguish between different research project archetypes. The first is 'moonshot'. This refers to one large and indivisible research project (think Manhattan Project or Moon Landings) which requires a large capital outlay upfront, or amortised sequence of payments. The return from the research project is only realised upon successful completion of the entire project.

The second is 'big pharma'. This refers to a sequential research project made up of individual stages, where commencing each stage is dependent upon the success of the previous one. Like the moonshot

² In Chapter 5, I explain how **research projects** are modelled as Type II endogenous uncertainties, where the probability of a **research project** succeeding is exogenous, but the timing of the observation of the project's success is endogenous, dependent upon the decision to start the project

archetype, the return from the project is only realised at the end of the process, in this case the successful completion of the final stage of the research project.

The final archetype is 'partial revenues'. As in the previous case of 'big pharma', this is a research project made up of sequential stages. However, in this case, the return from the project is realised incrementally as stages are successfully completed.

Highlighted technical terms and concepts are listed in the Glossary on page 229.

1.2 THE UK AS A CASE STUDY: POLICIES AND UNCERTAINTY

Revisiting the *energy trilemma*, the *sustainable* dimension was well served in 2008 when the Climate Change Act was enacted by the UK Government. This obligates the Secretary of State to reduce UK greenhouse gas emissions by 80% or more below 1990 levels by 2050 (HM Government, 2008). Setting such a long-term target, while necessary to counter the potential for stranded assets, due to the long lead-times and path-dependency associated with large energy investments, also serves to reduce one source of uncertainty regarding future carbon emissions: the political uncertainty associated with the UK's emissions budget. The carbon target is a good example of a *dynamic uncertainty*, where a pro-active decision reduces an uncertainty³.

However, for many other uncertainties to the energy system, for example, gas and oil prices, government regulation can only do so much to manage this static uncertainty⁴. Within energy system modelling, commodity prices belong to the class of *exogenous uncertainties*, where the uncertainties are outside the sphere of influence of the decision maker. The global market for oil and gas are subject to many of the same market influences as other commodities and some specific to these important energy resources. These include global economic growth boosting demand, changing socio-economic profiles of developing countries such as China and India whose burgeoning middle class increase the demand for private cars, and the geo-political com-

³ I avoid discussion of the social, ethical and economic aspects of greenhouse gas emissions to focus only on the perspective of uncertainty. However, if the UK's GHG emissions remained uncertain, then the risk of climate change falls upon the global population (including the UK's). There are many reasons to legislate an emissions target, but I am arguing that the fact that the target reduces uncertainty to many of the actors involved is often undervalued and overlooked.

⁴ This is a simplification. It is likely that uncertainty about commodity prices could be a hybrid of static and dynamic uncertainties. For example, a fraction of the uncertainty is epistemic i.e. due to our lack of knowledge and could be reduced through learning, whereas the remainder is inherent randomness of the global market system and beyond our ability to model. This touches on more philosophy of science aspects, such as questions of ergodicity, which are outside the scope of this thesis, although a small discussion is included in Appendix C

plexities of supply. It is difficult to see how UK policy makers could influence the uncertainty around oil and gas prices in the same way that it is possible in the previous example. Instead, strategies to manage these static uncertainties are required. This could include actions to reduce exposure to price volatility, shift away from oil and gas to alternative sources of energy, reduce demand by investing in energy efficiency, or secure indigenous resources.

Bringing such diverse uncertainties together into one policy or modelling framework seems challenging to say the least. As budgetary pressures are increased in the UK, analysis based upon [dynamic uncertainties](#) could reveal additional cost-saving measures which are explicitly linked to actions (investments). In contrast, studies based on a risk approach identify only defensive actions using technologies and under conditions about which we are currently certain.

1.3 UNCERTAINTIES AND THE ENERGY SYSTEM

There is only a relatively small body of work which has systematically investigated the effect of uncertainty upon the UK energy system using a modelling approach. One recent example using the [ETI-ESME](#) model is Pye et al., (2015). Energy system models have been mainly used to investigate the system-wide implications for the energy system under a range of long-term scenarios. A [UK Energy Research Centre \(UKERC\)](#) Energy Strategies Under Uncertainty workshop identified techniques and frameworks for managing uncertainties in the energy system (Davies et al., 2014). In particular, the study highlighted the ways in which decision makers understand and manage uncertainty in practice. The working paper highlighted that while there is a good deal of work on what is uncertain, techniques which accommodate learning (actions which resolve or reduce uncertainty) were not well understood among the participants in the workshop. As shown in Section 2.1.1, little research has focused on formally assessing which of the uncertainties are of most importance to the energy system. However, investigating how investments in uncertainty reduction, such as R&D, learning, or waiting, could help reduce the cost of the transition to a low carbon future has a larger presence in the literature. Learning is implemented into energy system models as an exogenous improvement in technology parameters (costs reducing and efficiencies increasing over time). Relaxing the assumption that learning is exogenous to the decisions, could highlight opportunities for investment that would otherwise be ignored.

There is potential for modified energy system models to directly inform the disbursement of research funds through quantitative modelling of the expected returns. The current methods for directing research funds revolve around strategic insight, peer review and the directorship of the UK Research Councils (<http://www.rcuk.ac.uk>).

However, (Low Carbon Innovation Coordination Group, 2014) shows how **ESOMs** have been used to inform the system aspects of the UK's energy innovation framework. Technology Innovation Needs Assessments (TINAs) have been compiled for the core technologies identified by industry and academia, including research using **ESOMs**. But there is little evidence of the final step of modelling the relationship between investment in energy innovation and the energy system, which is the primary contribution made by this thesis. In addition, the **Engineering and Physical Sciences Research Council (EPSRC)**'s strategy documents and the **Research Councils of the United Kingdom (RCUK)** Energy impact documents demonstrate that a range of considerations, other than the direct economic benefit, are taken into account when selecting **research projects**. For example, it is difficult to quantify the range of benefits of research, particularly as these are not only economic in nature. **RCUK** considers the *impact* of research in economic, societal and academic contexts, all aspects which are challenging to incorporate into models of the research process.

1.4 IDENTIFYING ACTIONS UNDER UNCERTAINTY

One benefit of distinguishing between static and **dynamic uncertainties** is that actions to reduce uncertainty are identified as part of the analysis. Static risk-based approaches to managing uncertainty are rarely an appropriate methodological response for situations in which it is difficult to quantify uncertainty using probability distributions⁵. A risk-based approach is only appropriate when good knowledge is available regarding both possibilities (what could happen) and probabilities (the likelihood of that event happening). However, for many of the inputs to an **ESOM**, our knowledge of probabilities of well-defined events is often limited. Scenario methods and sensitivity testing comprise two of the appropriate responses.

When the system aspects of the energy-economic-environment nexus are considered, the case against the use of risk-based approaches is compounded. For example, risk-based approaches may well be appropriate for the operational and engineering aspects of energy technologies in isolation. But when coupled with the **Knightian uncertainty** of economic and demographic projections, or the **ambiguity** surrounding environmental consequences of climate change at the local level, or unassailable **ignorance** of a radical, but unknown technology, it is not at all clear which, if any, methodological approach is most appropriate.

Another issue concerns the static nature of the existing uncertainty frameworks. These frameworks, while effective at discriminating between

⁵ See the theoretical framework of Stirling, (2007, 2010), reproduced on page 206, in which we can distinguish between two axes of uncertainty; knowledge about possibilities and knowledge about probabilities

different types of uncertainty, do not give insights into what to do when uncertainties change over time. Nor do existing frameworks indicate under which circumstances the actions of decision-makers can influence uncertainty. In other words, in addition to the fact that risk-based frameworks are not suitable for use across the whole system, the alternatives do not allow us to quantify the value of learning about energy system uncertainties. In the context of these challenges, it is clear that there is space for new tools which acknowledge the opportunities provided by dynamic uncertainties.

In designing new tools, there are a number of considerations to address. The first of which is relevance to policy. The outputs of a modelling framework, need to inform policy. For example, in the energy sector, there are a suite of policy tools available to decision makers. These include market-based instruments, such as subsidies & taxes, renewable portfolio standards, regulations and price support mechanisms. Many of these have the effect of providing support to technologies close to commercial competitiveness, when the majority of the uncertainty has already been resolved. Policies based on prices and regulations are relatively easy to model. In comparison, policies involving [dynamic uncertainties](#), such as R&D, are much more difficult to implement in existing models or decision support system. Investments in R&D programmes are used to support technologies and resources further away from the marketplace, when the majority of uncertainty is still to be resolved.

1.5 COMBINING A SYSTEM APPROACH WITH DYNAMIC UNCERTAINTY

Techniques which emphasise the role of decision makers in managing uncertainty have been popularised under the framework of real-options. However, the scale of energy system models present a prohibitive computational challenge to real-option approaches, which are typically solved analytically (although Monte Carlo approaches have also been used). As revealed in Chapter 2, real-options techniques are only one of a number of methods for dealing with [dynamic uncertainty](#), but there has been little progress in integrating [dynamic uncertainty](#) in ESOMs.

In ESOMs, currently no distinction is made between energy system uncertainties that are exogenous, endogenous, dynamic or static. There is no clear guidance as to whether taking a [dynamic uncertainty](#) approach would be beneficial, and no research currently available⁶ that have applied [dynamic uncertainties](#) to ESOMs. This potentially hampers the policy debate, by focusing on the management of

⁶ A very recent paper, Santen et al., 2016, applies a stochastic dynamic programming approach to RD&D investment in a single technology, solar. This utilises the same concepts of dynamic uncertainty and learning as this thesis.

risks, rather than the opportunities available to learn. However, techniques and methods for modelling dynamic uncertainty have been developed and are outlined in section 2.3.

In this study, the distinction between static and [dynamic uncertainties](#) in the [ESOM](#) are made explicit. The computational problem of dimensionality is circumvented through the use of a [global sensitivity analysis](#) technique to identify the most influential uncertainties. And a novel technique is applied to model the timing and choice of investments in learning while taking into account the systemic benefits of the uncertain results of those investments.

1.6 OVERVIEW OF THIS STUDY

The overall aim of this thesis is to link concrete actions with learning about uncertainties in the energy system. The traditional static view of uncertainty — as a fog occluding our view into the future — is paralysing in that it forces a defensive stance; one can only hedge against damaging outcomes⁷. Recognising that some uncertainties are epistemic, are dynamic, and amenable to learning presents options and actions for the present which move beyond the defensive. Chiefly, it is possible to invest in observing the true uncertainty of an outcome in addition to or instead of hedging against a costly outcome. This maybe of particular benefit for the energy sector in which cheap near-term investments are foregone due to the need for more expensive investments which provide a hedge against an expensive, but distant event. To express this in the language of a risk-manager; existing methods view the distribution of costs as fixed. This thesis presents an approach which allows the true value of an uncertain parameter to be observed through investing in research and development.

Of course, all static uncertainty does not disappear as soon as we change our perspective. Clearly, there are some uncertainties about which we cannot learn more (this is the definition of [aleatory uncertainty](#)). However, ignoring [dynamic uncertainties](#) reduces our options for action and at present, there is little research to quantify the magnitude of this effect.

Before distinguishing between static and dynamic uncertainties, it is first necessary to determine which uncertainties (of either type) are important. A method is needed to identify, rank or quantify the influence of energy system uncertainties. Then, a way in which dynamic and static uncertainties can be distinguished is needed. Identifying actions to link to the most important uncertainties is then required. Finally, evaluating these actions will give an insight into how these actions relate to the uncertainties which underpin the energy trilemma.

⁷ If the hedge is more expensive than the risk of the uncertain future event then there will be no perceivable difference between a hedge and the action ignoring the risk

1.6.1 *Research Questions and Objectives*

Four research questions arise from need to address the research problem discussed above:

1. What is the most appropriate [global sensitivity analysis](#) method to identify and rank uncertainties in energy system models, given the nature, source and type of [parametric uncertainty](#)?
2. What is the value of learning about critical energy system uncertainties?
3. What is the structure of dynamic energy-system uncertainties?
4. How do dynamic energy-system uncertainties relate to actions for learning or reducing these uncertainties?

The first research question highlights the importance of identifying a technique which is applicable to the specific case of energy system models. As discussed in Sections [2.1](#) and [2.2](#), the nature of the uncertainties treated by [scenario analysis](#) are quite different to the settings in which uncertainty is quantified using probability distributions - a situation which easily lends itself to a number of approaches.

The second question concerns the importance of quantifying the value of key uncertainties to the energy system. However, this does not specify what actions we should take given this information.

The third question relates to a framework for categorising dynamic uncertainties. In moving from the quantified value, in answer to the third question, and the actions of the final question, the uncertainties included in the analysis must adhere to criteria raised by this question.

The fourth question raises the issue of linking actions to uncertainties, an unavoidable consequence of examining and modelling uncertainties which change over time. In particular, what policy insights can we gain from linking research and development as one category action, to the resolution of critical energy system uncertainties?

The following research objectives reflect the steps that are required to answer the research questions:

1. Survey the available techniques for sensitivity analysis, match these candidate techniques to [ESOMs](#) and select an appropriate method (Chapter [3](#))
2. Perform a sensitivity analysis on the [ESME](#) model to identify critical energy system uncertainties for the model (Chapter [4](#))
3. Formulate and develop a learning model capable of evaluating the choice and timing of a learning process related to the critical energy system uncertainties (Chapter [5](#))

4. Select a subset of the critical uncertainties, applying an uncertainty framework that discriminates between different types of uncertainty, and evaluate their effect over discrete levels of availability/cost/efficiency etc. (Chapter 6)
5. Apply the learning model over a range of conditions to identify strategies and opportunities for learning about these critical energy system uncertainties (Chapter 6)

1.6.2 Academic Contribution

A review of energy system modelling studies reveals that despite a strong awareness of the implications of [structural uncertainties](#) and [parametric uncertainties](#), few systematic studies that focus on the behaviour of these models under uncertainty have been conducted. Of those studies which have been made, none have used a [global sensitivity analysis](#) approach to understand the relationship of the model outputs to the inputs over the entire model input space⁸. There seems to be no valid reason for a [global sensitivity analysis](#) not to be performed, other than the computational challenge of doing so. In other fields which use models of comparable size and complexity, such as integrated assessment modelling (see Section 2.2.3) [global sensitivity analyses](#) are used during model development and to support model insights.

Matching the attributes of the energy system model [ESME](#) with the available methods for conducting a [global sensitivity analysis](#), indicated that the Method of Morris is most suitable. This choice takes into account the nature of the uncertain inputs, data availability and computational demands. Alternative methods produce more detailed results, but at steep computational cost and require data at a resolution that is beyond what can be supplied for most [ESOMs](#).

Running the [global sensitivity analysis](#) for 119 (of 200) input parameters covering technology costs, efficiencies, resource costs, build rate constraints and resource constraints, the analysis shows that a remarkably small number of inputs are responsible for the majority of variation in the output of the [ESME](#) model. The same inputs appear across the sensitivity results from a range of candidate model outputs. These critical uncertainties include key energy resources, and systemic low-carbon resources and technologies.

A learning model linked to the [ESME](#) model provides a system-wide evaluation of these critical uncertainties, and determines strategies for learning about uncertainties. The strategies show how the scheduling of [research projects](#), can help reduce the cost of decarbonising the

⁸ Note that the term ‘global’ refers to a model’s input parameter space, rather than to the geographical boundaries of the model

UK energy system and relate the value of projects to the cost reduction in resolving an uncertainty.

The results show how the structure of the portfolio of learning projects strongly affects the expected profit from taking action. Dividing projects into discrete *stages* improves the financial prospects of the portfolio, particularly if the benefit of the project is only realised after the final *stage*. Front-loading of uncertainty is particularly beneficial, if it is possible to delay a more expensive, but more certain later *stage*.

1.6.3 *Thesis Outline*

The thesis is divided into seven chapters. After this introductory chapter, Chapter 2 introduces the relevant literature to the topics covered in the thesis. In the first section, the ways in which uncertainty has been managed in *ESOMs* are introduced. Sensitivity analysis, dynamic uncertainty and learning are then introduced in the remaining sections.

Chapter 3 introduces the *ESOM* called *ESME*, identifies and discusses the methods with which a *sensitivity analysis* can be conducted and proposes a novel combinatorial optimisation feature to improve the sampling method. Chapter 4 presents the results from the sensitivity analysis of the *ESME* model culminating in a brief discussion of the implications for energy modelling.

Chapters 5 and 6 present the learning model and the results. After a discussion of dynamic uncertainty in the context of modelling, the formulation of the learning model is then shown and a series of simple examples are used to demonstrate the explorative power of the selected framework. Chapter 6 present the results from the learning model. First, I explain the case study of critical uncertain parameters identified by the sensitivity analysis in Chapter 4. Then, the results from the learning model are shown from different perspectives. Understanding the relationship between the different strategies available to the learning model is interesting in itself. As such, a Monte Carlo sample is propagated through the model to explore the multi-dimensional space of the inputs and outputs. Finally, results from *ESME* are visualised over a tree which represents the possible outcomes from the combination of learning processes (*research projects*) associated with the availability of domestic biomass, and the build rate of CCS technologies. To place a value upon learning itself as a means to achieve a sustainable low-carbon future, the chapter identifies the plausible bounds of cost under which it would be considered effective to invest.

The final Chapter 7 concludes the thesis, summarises the key points from each of the chapters and makes suggestions for future work based on the findings from the work contained within. An Appendix to the thesis contains a glossary, a full bibliography and additional supporting material that is referenced throughout the text. Highlighted

words refer to terms that are defined in the Glossary on page [229](#) onwards. In the electronic version of this document, clicking on the highlighted word will activate a hyperlink to the relevant entry in the glossary. References are also hyperlinked.

This chapter is divided into five parts. In the first section, I explore the ways in which energy system modellers have identified parametric and [structural uncertainties](#), and the use of [scenario analysis](#) and Monte Carlo techniques for managing uncertainty. In the second section, I investigate the available methods for conducting sensitivity analyses to identify and rank key uncertainties. I then explore which of these methods have been applied to [Integrated Assessment Models \(IAMs\)](#) and [Energy System Optimisation Models \(ESOMs\)](#). In the third section, I look at various techniques for modelling [dynamic uncertainty](#) including [stochastic programming](#) and [stochastic dynamic programming \(SDP\)](#). In the fourth section, I investigate the literature related to learning in the energy system. While the literature in the fourth section is largely situated in the context of [IAMs](#), the concepts are closely related to those of [dynamic uncertainty](#). A final section connects these four parts of the literature together with links to the research aims and objectives outlined in Chapter 1, and the subsequent method and results chapters. This section defines a number of research questions.

2.1 TREATMENT OF UNCERTAINTY IN ENERGY SYSTEM MODELING

The scope of the literature view is focused in the following ways. I refer predominantly to [ESOMs](#). These are models which compute an optimal, generally least-cost, energy system, using a linear programming formulation in their fundamental form. A technical description of [ESOMs](#) can be found in Section 3.1 in which I introduce the [ETI-ESME](#) model. The aim of this part of the review is to determine how this field has incorporated an understanding of uncertainty and the methods that have been used to manage uncertainty within these models.

In comparison to other disciplines, such as environmental science, the interdisciplinary [ESOM](#) field has been relatively slow to adopt approaches that offer a comprehensive characterisation of uncertainty. Typically, [ESOMs](#) are run using a scenario approach, with a small set of central, high and low scenarios (Section 2.1.2). Typically, large numbers of sensitivity scenarios are run around each of the initial set of scenarios (Rosenberg et al., 2010), essentially a limited one-at-a-time sensitivity analysis as outlined in Saltelli et al., (2010a). Selected [ESOM](#) models incorporate additional techniques for managing uncer-

tainty, including stochastic optimisation on key epistemic uncertainties (Usher et al., 2012) (also see Section 2.3.1). However, ESOMs are typified by their complexity, large numbers of uncertain and correlated inputs, lack of empirical underpinning of inputs, and finite modelling resource for uncertainty analysis (Usher et al., 2013). Defining an uncertainty model for the inputs to ESOMs is particularly challenging. Energy system models rely upon forecasts and projections of key input parameters, such as GDP, population and demographic trends, many of which are not available in probabilistic form. And numerous forecasts and projections are themselves the output of other deterministic models.

2.1.1 *Parametric and Structural Uncertainty in ESOMs*

Broadly, there are two sources of uncertainty in models (Morgan et al., 1992; Spiegelhalter et al., 2011a; Wynne, 1992). The first is *parametric uncertainty* - that included in the definition of the inputs (or parameters) of the model. The second is *structural uncertainty* - uncertainty in the mathematical formulation of the model. Both *parametric uncertainties* and *structural uncertainties* are acknowledged in ESOM studies, although the former is investigated more readily than the latter.

2.1.1.1 *Parametric Uncertainty*

The original response to uncertainty in the parameters of future technologies was ‘simplicity’ (Fishbone et al., 1981) - in that the costs of technologies were exogenously specified, rather than endogenised within a technology learning formulation. Thus uncertainty was acknowledged as parametric, rather than structural, in form and could be investigated through sensitivity studies. Rath-Nagel, (1982) discuss ESOMs in the context of reducing oil dependence, with ‘planners’ responsible for the allocation of resources to fund R&D. The model generator used, MARKAL in this case, was intended as a tool for constructing alternative perspectives on the future of energy, while understanding the results within the context of explicit assumptions. The benefits of the model include the detail through the upstream and supply chain, so that it is a

“... a straightforward matter to assess the relative competitiveness of technologies in response to changes in important parameters and availability of fuels (Rath-Nagel, 1982).”

However, little investigation is conducted into identifying exactly which parameters are important, or quantifying how important they are.

In the decade that followed, stochastic programming formulations of the MESSAGE and MARKAL models were used to investigate un-

certainties in technology costs and abatement policies (Condevaux-Lanloy et al., 1996; Messner et al., 1996). A series of papers used stochastic programming to investigate uncertain energy demand and mitigation efforts (Kanudia, 1998; Kanudia et al., 1999; Loulou et al., 1999). Hu et al., (2010a) explicitly model policy uncertainty using a stochastic variant of MARKAL, and Usher et al., (2012) use stochastic programming to investigate critical mid-term uncertainties for decarbonisation of the UK energy system. In each of these studies, the representation of parametric uncertainty is limited to very few, or individual uncertainties. Formal methods, such as sensitivity analysis, to identify which uncertainties are most important for the model solution are notable in their absence. However, these papers do introduce the concept of strategies which hedge against future uncertainties. As will be discussed in Section 2.3, these techniques treat uncertainty as something that is exogenous to the decisions.

By far the most common methodology for running ESOMs, such as MARKAL and MESSAGE, is through the use of scenarios. The use of scenario analysis is appropriate given the nature of many of uncertainties in energy system models. More detail is given in Section 2.1.2 below.

Uncertainty in the parameters of energy system models is frequently addressed as an important concern of the authors of journal papers. For example, Anandarajah et al., (2010), make the following observation:

“In any energy model, a range of key parameters have deep uncertainties (especially in the longer term) and corresponding major impacts on energy systems pathways and costs.”

However, there is no consistent understanding of what uncertainty means between authors.

Within the ESOM literature, there are pockets of insight into the role of parametric uncertainty within these models. Golodnikov et al., (1995) illustrate the sensitivity of electricity production from one technology, CCGT plants, to the change in capital costs of competing plants. The same observation, that there are important interactions between input assumptions, is made by Rosenberg et al., (2010), who notes that it is the relative differences between fossil fuel prices that are most important in determining optimal technology selection, rather than absolute prices. In the former case, this observation is used as justification for using a stochastic programming approach to avoid the penny-switching behaviour of an optimisation model. One of the more comprehensive studies is by Yeh et al., (2006) who perform a local parametric sensitivity analysis of the transport sector to investigate the future uptake of hydrogen transport.

The scenario analysis presented in Rafaj et al., (2007) highlights some key uncertainties. These include: valuing socio-political priorit-

ies of future energy sector developments, socio-political acceptance of technological options, income distribution effects, discounting of the future damages to the present value, regional differences in valuing externalities, rate of technological change. However, all of these are not explicitly parameterised within the energy system model. Rather, combinations of parameters within the scenario framework are used to model many of these high-level drivers.

2.1.1.2 Structural Uncertainty

Kann et al., (2000) discuss model comparison as a means to understand the structural differences between models, as well as the differences in mathematical formulation that come about due to design decisions which flow from subjective value judgements about how to distill real-world concepts into a mathematical model. The use of model comparison to investigate uncertainty spans large projects or different flavours of the same model, as in Chen et al., (2007) who compares three different MARKAL variants - elastic demand, a standard version, and a macro-economic linkage, all for mainland China. Bosetti et al., (2015) harmonised inputs across three integrated assessment models, investigated the sensitivity of the models to technology uncertainties, and compared the results from each of the models to establish the importance of structural assumptions across the time horizon, output metrics and climate policy stringency. The EMF conduct regular model comparison exercises using a range of IAMs such as Fawcett et al., (2009). The typical approach is to develop a set of scenarios, which are then modelled independently by research groups around the world. Results are then compiled into a standardised template, written up into journal papers, where the model results, and models themselves are contrasted and compared with one another. The model comparison projects meet challenges when the partner models are unable to harmonise inputs, due to fundamental structural differences, or are unable to produce results due to infeasible areas of the models input and output spaces (an example can be found in Strachan et al., 2008c). These areas of missing data can make it difficult to obtain useful insights across the range of models. On the other hand, the exercise can be extremely useful for individual teams who stretch their model beyond its normal zone of use. One potential criticism of model comparison exercises is the tendency for sanitised average results to be reported, which mask the disagreements or structural differences between models (Socolow, 2011).

Table 1: Energy System Modelling Studies 1979 to 2012

Citation	Model	Technique
Agnew et al., (1979)	MESSAGE I	

Table 1: Energy System Modelling Studies 1979 to 2012 continued

Citation	Model	Technique
Avenhaus et al., (1980)		
Fishbone et al., (1981)	MARKAL (BNL)	OAT
Schrattenholzer, (1981)		
Nurminski et al., (1982)	MESSAGE II	
Rath-Nagel, (1982)	MARKAL	SA
Berger, (1990)	MARKAL-Ontario	S
Larsson et al., (1993)	MARKAL	MH
Larsson, (1993)	MARKAL	MH
Condevaux-Lanloy et al., (1996)		
Messner et al., (1996)	MESSAGE III	SP
Wene, (1996)	MESSAGE II, ETA-MACRO	ML
Kanudia, (1998)	MARKAL India	SP
Kanudia et al., (1999)	MARKAL	SP; MR
Loulou et al., (1999)	MARKAL Quebec	SP; MR
Seebregts et al., (1999)	MESSAGE, MARKAL, ERIS	M
Condevaux-Lanloy et al., (2000)	MARKAL	SP
Salvia et al., (2002)	MARKAL	SS
Barreto et al., (2004)	GMM	SS
Chen, (2005)	China MARKAL-MACRO	S
Tseng et al., (2005)	US MARKAL	S
Endo et al., (2006)	MARKAL	S
Ichinohe et al., (2006)	MARKAL	SA
Yeh et al., (2006)	US EPA MARKAL	SA, S, MC
Chen et al., (2007)	China MARKAL	M
Contaldi et al., (2007)	MARKAL-MACRO	S
Das et al., (2007)	Various Models	M
Endo, (2007)	MARKAL Japan	S
Krzyzanowski et al., (2007)	GMM	SS
Rafaj et al., (2007)	GMM	S
Schulz et al., (2007)	Swiss MARKAL	S;SA
Contaldi et al., (2008)	MARKAL	SS
Jiang et al., (2008)	MARKAL China	SS
Labriet et al., (2008)	ETSAP-TIAM	SP
Strachan et al., (2008e)	UK MARKAL-Macro	SS
Strachan et al., (2008a)	UK MARKAL-Macro	SS
Strachan et al., (2008d)	UK MARKAL-Macro	SS
Yeh et al., (2008)	US EPA MARKAL	S
Contreras et al., (2009)	MARKAL	S
Gül et al., (2009)	GMM	S

Table 1: Energy System Modelling Studies 1979 to 2012 continued

Citation	Model	Technique
Held et al., (2009)	MIND	CP
Kannan et al., (2009)	UK MARKAL	S;M
Kannan, (2009)	UK MARKAL	S;OAT
Loulou et al., (2009)	ETSAP-TIAM	SP
Strachan et al., (2009)	UK MARKAL	SS
Broek et al., (2009)	MARKAL	CBA
Akimoto et al., (2010)	DNE21+	S
Anandarajah et al., (2010)	UK MARKAL	S
Heaton et al., (2010)	ESME	MC
Hu et al., (2010b)	MARKAL	SP
Labriet et al., (2010)	ETSAP-TIAM	SP
Rosenberg et al., (2010)	MARKAL	SS
Strachan, (2010)	UK MARKAL	S
Broek et al., (2010)	MARKAL	S
Wright et al., (2010)	MARKAL	SS
Anadon et al., 2011	MARKAL	MC
Babonneau et al., (2011b)	TIAM-WORLD	RO;CP
Babonneau et al., (2011a)	GEM-E3;TIAM-WORLD	SP;MC;ML
Baker et al., (2011c)	GCAM	VOI
Beugin et al., (2011)	CIMS	MC;B
Bréchet et al., (2011)	S-CWS	SP
Cian et al., (2011)	WITCH	SP
Keppo et al., (2011)	TIAM-ECN	SP
Lorenz et al., (2011)	MIND	SP
Strachan et al., (2011)	UK MARKAL	S
Usher et al., (2012)	UK MARKAL	SP
DeCarolis et al., (2012)	TEMOA	SP

OAT one-at-a-time local sensitivity analysis ^{SA} local sensitivity analysis

^{SP} stochastic programming (two stage or multi-stage)

^{SS} scenario analysis with sensitivity scenarios ^S scenario analysis only

^{MC} Monte Carlo sampling ^M model comparison

^{MR} minimax regret ^{MH} manual hedging ^{ML} model linkages

^{CP} chance-constrained programming ^{CBA} cost-benefit analysis

^{VOI} value of information ^B backcasting

As shown in Table 1, of the sample of 80 or so studies published up to 2012 in peer reviewed journals that have used energy system models as a core component of their analysis, very few have comprehensively adopted uncertainty as a central tenet of their investigation. Over time, there has been a trend towards running more and

more scenarios using models that are increasingly complicated and detailed.

2.1.2 *Formal Scenario Analysis*

Typically, energy system modelling and to a large extent, integrated assessment modelling has focused on a scenario type approach to manage future uncertainties. These uncertainties are inevitable when dealing with the 50 year or more time-horizon needed when investigating energy-economic transitions. In this section, I present a brief history of scenarios, followed by a discussion in which I compare the aspirations of those who invented and popularised the approach, with the evidence from the body of scenario literature on [ESOMs](#).

Scenario analysis commenced in the 1950s with the publication of Kahn's controversial book 'On Thermonuclear War' (Kahn, 1960) who investigated the influence of direct nuclear weapon strikes upon US cities. Scenario analysis was subsequently used for military and strategic purposes, and then popularised by the oil giant Shell in the 1970s and 1980s (Wack, 1985a,b) after Shell's success in navigating the oil crises relatively unscathed, a feat largely put down to their management's use of scenarios. Scenarios largely grew in popularity as an alternative to the various techniques used for forecasting (Huss, 1988). The recognition of the chronic failure of forecasts and projections during this period produced a demand for new techniques such as the use of scenarios and techniques for judging robustness (Shlyakhter et al., 1994; Smil, 2000). Perhaps the most visible use of scenarios in the energy and climate sphere recently has been the development of the emission scenarios by IPCC, (2000), and subsequent [Representative Concentration Pathway \(RCP\)s](#) (Meinshausen et al., 2011).

Miller, (2003) identify an eight step process for the development of scenarios, commencing with the definition of a scope; the elicitation of expert knowledge regarding what is known; the identification of a destination; the sketching of possible paths to a future point; identification of uncertainties; tests for plausibility; and finally the formulation of strategies to deal with identified interactive dynamics and uncertainties. To paraphrase, scenarios, by acknowledging, structuring and understanding uncertainty create a small selection of alternative and internally consistent pathways into the future.

As developed by Shell, scenarios were intended:

“...to help decision makers develop their own feel for the nature of the system, the forces at work within it, the uncertainties that underlie the alternative scenarios, and the concepts useful for interpreting key data...” (Wack, 1985a)

[Scenario analysis](#) has remained the primary underpinning technique of energy system modelling, and the main method by which

uncertainties are managed. However, a large number of studies contravene the best-practice outlined by Wack, (1985a), particularly that scenarios should never number more than four, and that three scenarios should be avoided for fear of misinterpreting the middle scenario as a central estimate (see for example Anandarajah et al., 2010). This is not to suggest that such studies are poor, but to indicate that the computational and data management techniques available today are far in excess of what was available when *scenario analysis* was first proposed.

A key tenet of *scenario analysis* is the use of ‘internally consistent’ input assumptions. This has also proved increasingly problematic as the size of models increases in accordance with available computing power. One potential solution to this is the use of *Bayesian Networks (BNs)* to quantify the probabilistic relationship of parameters with one another. Cinar et al., (2010) applied a *BN* to generate a causal probabilistic network for the Turkish energy system. By eliciting expert beliefs of the structure of the parameters, Cinar et al., (2010) then used historical data to assess the probabilities associated with changes in the different parameters. An advantage of this approach is that the input assumptions are consistent with one another across a range of scenarios, while causal factors are represented explicitly. Usher et al., (2013) conducted an expert elicitation of UK energy experts to quantify uncertainty around a selection of drivers of energy consumption. The results showed that experts agreed more strongly on a dependence structure between uncertain parameters than the distributions of the uncertain parameters.

Trutnevyte et al., (2016) review UK scenarios developed between 1978 and 2002. They find that the choice of scenarios was largely dictated by contemporary debates at the time of the analysis. They claim that scenarios tend to focus on *parametric uncertainties*, rather than structural changes that are harder to *parameterise*, such as changes in governance, shifts in environmental concerns and restructuring of industry.

2.1.3 *Monte Carlo Simulation*

Another method that has gained popularity in the field of energy system modelling is *Monte Carlo* sampling. A Monte Carlo experiment is a way of propagating uncertainty through a model by sampling the joint-distribution of inputs to a model many times, running the model for each of these samples, and observing the distribution of the output(s). The *ETI-ESME* model is set up to run in this mode (as demonstrated in Pye et al., 2015), but the technique has some methodological issues when applied to *ESOMs*. For example, for many of the uncertain parameters, it is difficult to quantify the uncertainty using a probability distribution as there is little or no data available.

This means that it is not feasible to assign objective probability distributions. The use of subjective beliefs to quantify uncertainty, for example using expert elicitation (Anadon et al., 2014; Usher et al., 2013) is time consuming. Directly eliciting correlations between uncertain parameters is extremely difficult, even for experts (O'Hagan, 2006a). Eliciting proxies for correlation is more manageable, for example using Spearman Rank Transformations (Anadon et al., 2011).

Interpreting the output from a Monte Carlo simulation of an ESOM is also problematic. Given that an ESOM computes an optimal solution within a deterministic framework, running the model multiple times with different input values results in pathways which may contradict one another. For example, a range of values for a key uncertain parameter may give divergent near-term investment pathways. Following any individual pathway locks out the other pathways.

2.2 SENSITIVITY AND UNCERTAINTY ANALYSIS

In this section, I give an overview of sensitivity analysis. I touch on the very few previous studies within the ESOM literature which have made use of some of these techniques. Scenario analysis has been the predominant technique for managing uncertainty for ESOMs, so the studies which have used sensitivity analysis are in the minority. As such, I draw upon work from the IAM literature which have made greater use of sensitivity analysis techniques. A detailed technical discussion of one sensitivity analysis technique which is used heavily in this thesis, the Method of Morris, is covered in Chapter 3.

Before investigating sensitivity analysis in greater detail, it is worth considering the ways in which the uncertainty surrounding parameters may be quantified. A number of studies which have examined the techniques for eliciting expert judgements (Morgan et al., 1992; Morris, 1977; O'Hagan, 1998, 2006b; Tetlock, 2005), and communicating uncertainty, Spiegelhalter et al., 2011b. Some of these techniques have subsequently been used in an energy setting (Anadon et al., 2012; Baker et al., 2009, 2010, 2011a; Usher et al., 2013). Given that many of the parametric and structural uncertainties associated with ESOMs are not measurable in the conventional sense, expert judgement is one of the main methods for producing measures of uncertainty, such as probability distributions, which can then be used in numerical models. Keith, (1996) outlines three options for combining the beliefs of multiple experts, often required when developing ESOMs which incorporate many uncertain parameters. They include the Delphi method, which is susceptible to bias; an averaging method (Kaplan, 2000; Keith, 1996; Morris, 1977), which makes implicit or explicit assumptions regarding the weight of the experts (how expert is each expert?); or guessing! For example, the probabilistic data used in the ETI-ESME model was arrived at through a consultation pro-

cess with stakeholders, which approximated the Delphi method. In Chapter 6, I populate the learning model with “made-up” probabilities for demonstration purposes. In all cases, it is worth examining the underlying assumptions which lead to any probability distribution, just as it is good practice to test any other assumption. The tools outlined in this chapter enable powerful and systematic investigation of these assumptions.

2.2.1 *Sensitivity Analysis is not Uncertainty Analysis*

Global [sensitivity analysis](#) is a family of techniques used to determine the influence of a model input upon a model output. [Sensitivity analysis](#) is often conflated with [Uncertainty Analysis \(UA\)](#) and [uncertainty propagation \(UP\)](#), but these are both distinct steps within a full meta-study of a model’s response to uncertainty. An [UA](#) is the process of quantifying the uncertainties in a model’s inputs through expert elicitation, statistical analyses and so on. Then, the aim of [UP](#) is to quantify to what extent uncertainty exists in the outputs of a model by using a technique, such as Monte Carlo sampling (or alternative), to propagate uncertainty through a model. However, the aim of sensitivity analysis is to, independently of the uncertainty of a parameter, determine each parameter’s influence upon the model output. So to paraphrase Hamby, (1994), an uncertainty analysis ranks parameters according to their importance, and a sensitivity analysis ranks parameters according to sensitivity. “

An important parameter is always sensitive because parameter variability will not appear in the output unless the model is sensitive to the input. A sensitive parameter, however, is not necessarily important because it may be known precisely, thereby having little variability to add to the output (Hamby, 1994).”

Morgan et al., (1992, ch.8) describe how using the two tools of [sensitivity analysis](#) and [UP](#) enables modellers to understand the relative importance of:

- model inputs
- structural assumptions
- consequences of modelled decisions, and
- disagreements over input values or choice of inputs

they list mitigating actions, including:

- whether to gather more information
- make more careful uncertainty assessments, or

- refine the model

Morgan et al., (1992) give an overview of the techniques for assessing model sensitivity. These include the assertion that the elementary unit of sensitivity is gradient, they identify the importance of global versus local [sensitivity analyses](#), and the chicken-and-egg situation around whether to do a full [UA](#) versus a more limited [sensitivity analysis](#) when little is known about a model's sensitivity to assumptions.

2.2.2 *Types of Sensitivity Analysis*

Sensitivity analyses are divisible into local and global approaches. A local approach, such as the one-at-a-time (OAT) approach, are conducted by increasing or decreasing the value of each parameter in turn, returning the parameter to its central value before moving onto the next. Local approaches ignore the potential for interactions between inputs, but are simple to perform. Computationally, the approach is inefficient, but often one-at-a-time variation of input parameters is performed manually for only a few variables. Another disadvantage is that interaction effects between input variables are ignored. Interaction effects are when a parameter becomes more or less influential depending upon the value of a one or more other parameters (Saltelli et al., 2008a).

Global sensitivity analyses are the state of art and give a much more robust measure of the influence of input parameters over the entire model input space. A huge number of model runs are required to fully explore the input model space which can quickly become computationally infeasible. Given continuous input distributions, the full input space can never be fully quantified, thus there is a risk that some interaction or effects are omitted from the analysis. However, if there is sufficient coverage of the input space, then statistical techniques can be applied to provide estimates of parameter influence upon the model outputs, and in most cases, bootstrapping can provide confidence intervals to ensure that the input sample is good enough to offer robust rankings of input parameters.

Saltelli et al., (2008a) outline the four settings within which a sensitivity analysis can be used. These are:

- factor fixing: screen out non-influential factors
- factor prioritisation: identify the most important factors
- factor mapping: identify the factors associated with a portion of the output space
- meta-modelling: build a statistical representation of the relationship between a model's inputs and outputs

In general, factor fixing and factor prioritisation are the most common settings.

Saltelli et al.'s published work advocates for the use of variance based approaches to computing sensitivity. Variance-based approaches build upon the fundamental idea that the output variance of a model output can be decomposed into increasing dimensions of arbitrary combinations of variables (Sobol', 1990; Wagner, 1995). Computation of indices which reflect the influence of the input variables follow. Although the techniques are similar, different authors have proposed alternative means of computing sensitivity indices. A focus of the research literature has been on the sampling methods — how to achieve robust estimates of sensitivity indices at least computational cost. Saltelli et al., (2010b) review methods for performing a global [sensitivity analysis](#) that gives a 'total sensitivity index' S_{T_i} , measuring the first and higher order effects (interactions) of an input factor.

Archer et al., (1997) compare the use of Sobol' sensitivity indices to the statistical method analysis of variance (ANOVA). While others came up with alternative methods of computing sensitivity indices using variance approaches, the authors state that Sobol''s is the most general, encompassing that of the Fourier amplitude sensitivity test (FAST) (Cukier et al., 1978). The similarity was then shown in Saltelli et al., (1998). Saltelli, (2002) extend Sobol''s original method so that it out performs the up-to-then more efficient FAST method of Cukier et al., (1978).

Morris, (1991), extended in Campolongo et al., (2007, 2011) propose a simple extension to one-at-a-time sensitivity analyses which gives a proxy for total sensitivity index. The method works by progressively incrementing individual parameter values over discrete [levels](#) from randomly generated starting [trajectories](#). Wainwright et al., (2014) compare the application to hydrological models of Morris and the Sobol'/Saltelli variance-based approach. They conclude that both methods provide complementary information. The Morris method is computationally cheap, and the metric μ^* (the mean of the elementary effects) is a good proxy for the total sensitivity index.

Oakley et al., (2004) offer a Bayesian perspective on sensitivity analysis, reviewing the variance-based approaches of Saltelli. Their approach offers new prospects for future performance improvements in sensitivity analysis (Saltelli et al., 2005). The approach infers the value of the main-effect and total-effect indices from just one set of model runs (i.e. without needing to rerun a model specifically for the [sensitivity analysis](#), instead making use of existing model runs). The approach does require an assumption of smoothness of the model output, and as such is inappropriate for models with highly non-linear outputs. The technique is very computationally efficient in comparison to other techniques as the smoothness assumption allows the approach to make use of the information available after each successive

simulation run to infer the value of the model output at nearby locations. In contrast, the variance-based approaches which use Monte Carlo sampling cannot use this information as each model run is treated independently from the others.

Plischke et al., (2013) identifies some drawbacks of variance-based approaches, namely the assumption that inputs are independent of one another. The authors show that in special cases where interaction effects are more important than main effects, variance-based approaches which do not explicitly take dependence-relations into account can fall foul of Type I errors (false positives). Kucherenko et al., (2012) extends the variance-based approaches using copulas to dependent inputs, but at high computational cost.

When a model is severely computationally demanding, a factorial or fractional factorial approach may give some initial insights within the factor fixing setting. A factorial approach is a two-level global approach, moving an input within two extremes of its range, or turning an input on and off. Saltelli et al., (1995, 2008a) show how using a Resolution IV design-of-experiments approach can be used to compute a main-effect which incorporates a measure of interactions with fewer models runs than the number of input factors.

Generally, there is a trade-off between the computational demands of the methods, the number of inputs they can handle, and the detail of the measures of sensitivity. Some methods allow the grouping of inputs which increases the coverage of inputs, and reduces the computational time at the expense of aggregated results. The different methods also have differing requirements for input data (from ranges to distributions), and capabilities for managing correlated inputs and dependence structures. There is also a division between the approaches which require the generation of sample, which is then run through the model to generate results which are then analysed in conjunction with the input sample, and techniques which post-analyse a given Monte Carlo run of a model. These attributes are summarised in Table 5 in Section 3.2.

2.2.3 *Applications of Sensitivity Analysis to Energy Models*

There are a number of energy related studies which have made use of formal sensitivity analysis approaches, as shown in Table 2. Particularly within the IAM community, [sensitivity analysis](#) has been used both for model dimensionality reduction (factor fixing) and to determine those parameters of most influence (factor prioritisation).

One of the earliest examples of the method of Morris used on an energy-economic model culminates in Sluijs et al., (2005) who perform a sensitivity analysis of the IMAGE/TIMER model using a revised version of the method of Morris. Campolongo et al., (1999) conducted the sensitivity analysis on three outputs and six inputs, requir-

Paper	Model	Method	Variables
Sluijs et al., (2002)	TIMER	Morris	300
Sluijs et al., (2005)	IMAGE, TIMER	Morris	
Campolongo et al., (1999)	TIMER	Morris	6
McJeon et al., (2011)	GCAM	full factorial	99
Scott et al., (2014)	GCAM	fractional factorial	5
Anderson et al., (2014)	DICE	Plischke et al., (2013)	51
Butler et al., (2014)	DICE	Sobol', (1990)	30
Bosetti et al., (2015)	WITCH, GCAM, MARKAL	Plischke et al., (2013)	8
Branger et al., (2015)	Res-IRF	(Morris, 1991)	71

Table 2: A summary of the global sensitivity analysis techniques applied in energy related studies

ing 128 simulations to obtain the results. The input ranges used were an arbitrary $\pm 10\%$, rather than the full plausible range over which the parameters could vary. Sluijs et al., (2002) performed a Morris analysis of TIMER, a component of the IMAGE model. In this case, the Morris analysis was used to screen out the non-influential parameters, reduce the dimensionality of the model, and focus uncertainty quantification on only the most important (influential) input parameters. The TIMER model has around 300 input parameters, with 160,000 data points comprising the input data for these parameters, many of which are time series representing historical and projected trends. The report suggests that choosing arbitrary bounds, in their case a $\pm 50\%$ range around the base values could have biased the results of the analysis.

McJeon et al., (2011) use a full factorial approach, and run 768 scenarios exploring combinations of technology assumptions against two stabilisation goals in the large integrated assessment model GCAM. Scott et al., (2014) then apply a fractional factorial approach to the USA version of GCAM. Their results show that carbon capture and sequestration technologies are important to reduce the cost of mitigation scenarios, particularly when technological innovation in the building and transport sectors is poor. The sensitivity analysis also shows that there are many possible combinations of technologies which do not differ radically in terms of cost, thus technologies can substitute for one another in many permutations. However, the choice

of fractional factorial approach is questionable because while interaction effects can be identified, non-linear effects cannot be identified with any approach that uses just two-levels (Saltelli et al., 2008a).

Anderson et al., (2014) and Butler et al., (2014) both perform global sensitivity analyses of the integrated assessment model *DICE*, the former using the technique of Plischke et al., (2013) which allows extraction of sensitivity results from the same Monte Carlo run used in a standard uncertainty analysis, the latter of Sobol', (1990).

Bosetti et al., (2015) use the techniques of Plischke et al., (2013) to compare the influence of eight technology attributes on various outputs across three models, *WITCH*, *GCAM* and *MARKAL*, after the performing an *UA* using the same Monte Carlo sample. The results reveal many structural insights into the differences between the models, but it is perhaps a pity that the study omitted running a truly global sensitivity analysis which perturbed parameters outside of the eight investigated.

Branger et al., (2015) conduct a sensitivity analysis using the method of Morris on an energy-economic model of French dwelling space-heating demand, first conducting an *UA*, followed by an *sensitivity analysis*. The input uncertainties are quantified using probability distributions, and the Method of Morris (Morris, 1991) is used in the settings of factor prioritisation and factor fixing.

2.2.4 Use of sensitivity analysis in other fields

The majority of the literature on sensitivity analysis is concerned with the types of models for which probabilistic inputs are used. For example, in the natural sciences, there is a strong literature on the use of variance-based global sensitivity analysis techniques to screen out unimportant input variables and identify the input factors of largest influence. For example, Herman et al., (2013) and Srivastava et al., (2014) apply global sensitivity analysis techniques to models of hydrological systems. Nossent et al., (2011) apply the techniques of Sobol', (1990) to a complex environmental model of a river system to perform factor prioritisation and fixing. Lagerwall et al., (2014) use Sobol' for a global sensitivity analysis of an ecological model. Ireland et al., (2015) is an example of the application of the Bayesian techniques proposed by Oakley et al., (2004) to a land-biosphere model.

Painter et al., (2007) compare the use of scenarios with one-at-time (OAT) sensitivity analysis, factorial and fractional factorial approaches. While simpler than the other approaches, the scenario method does not identify interactions between inputs, while the factorial and fractional factorial approaches do identify interactions. The latter approach is an efficient method of screening input parameters which identifies interactions between variables.

2.3 MODELLING APPROACHES WITH DYNAMIC UNCERTAINTY

In this section, I introduce the major works in modelling [dynamic uncertainty](#). The literature is divided into papers which use a stochastic programming methodology, and those that are derived from a [stochastic dynamic programming \(SDP\)](#) technique. I also include a brief section on stochastic decision trees. While stochastic programming is associated with operations research, stochastic dynamic programming is more common in economics and in engineering situations, such as optimal control. However, the modelling of [dynamic uncertainty](#) can be approached using either of the mathematical techniques and each is challenging in its own way.

Until now, I have used the term [dynamic uncertainty](#) to represent any situation in which uncertainty can change. However, in the modelling literature the term [decision-dependent uncertainty \(DDU\)](#) is often used. This extra term is useful to refer explicitly to uncertainty which changes as the direct result of a modelled decision. Within a model, uncertainty is parameterised using probability distributions (or equivalent such as a decision tree). The difficulty in modelling decision-dependent uncertainty arises because the uncertainty is affected by the outcome of a decision, or because the timing of the observation of an uncertainty is dependent upon the decision. The problems representing this are described in the subsections below.

This section avoids a more technical description of the particular challenges of modelling decision-dependent uncertainty as these details are introduced in the later methodology Chapter 5 and Section 5.1.

2.3.1 *Stochastic Programming*

In the stochastic programming version of the [dynamic uncertainty](#) problem, the representation of decision dependent uncertainty requires that the structure of [scenario tree](#) itself is dependent upon the value of the [decision variables](#). In other words, the full [scenario tree](#) branches at each decision and for each uncertain parameter. In the deterministic equivalent formulation of the problem, the [non-anticipativity constraints \(NACs\)](#) become conditional, and are imposed dependent upon the value of the decision variables. The advantage of the stochastic programming formulation is that the deterministic equivalent of this problem can be formulated as a mixed-integer program, for which powerful commercial algorithms are available. One potential disadvantage is that continuous dependent variables and uncertainties cannot be represented, as discrete variables are required to compose a [scenario tree](#).

Two further terms to define uncertainty are introduced in the stochastic programming literature focused on [dynamic uncertainty](#). [Exogenous](#)

uncertainty is that which is external to the model and cannot be affected by the **decision variables**. **Endogenous uncertainty** is that which can be affected by the **decision variables** and thus has the same meaning as decision-dependent. What is, or is not, an exogenous or endogenous uncertainty is a function of the design decisions made when formulating the model.

The PhD thesis of Jonsbraten, (1998) and the paper contained within (Jonsbraten et al., 1998) introduces a stochastic programming framework in which the distribution of random parameters is dependent upon the decision variables. The decisions which determine the realisation of uncertainty are constrained to the first period only, simplifying the formulation. This work was conducted in the context of decision support for the exploitation of an oil field under uncertainty, where information revealed about the nature of the oil fields is dependent upon the decision to drill. A series of methodological papers, all with applications in oil drilling, build upon and refine the work of Jonsbraten, (1998). The work of Goel et al., (2004) expands the approach to multiple gas fields. A later work (Goel et al., 2006) fully generalised the approach, extending the multi-stage stochastic programming framework to decision dependent uncertainty. Tarhan et al., (2009, 2011) and Gupta et al., (2011b) expand the work of Goel et al., (2006), by incorporating non-linear constraints to better represent oil fields, and treating uncertainty resolution as a gradual rather than instantaneous process. Gupta et al., (2011a) identify additional properties to that of Goel et al., (2006) to reduce the number of non-anticipativity constraints, thereby allowing the solution of larger problems. Gupta et al., (2014) develop an alternative technique for relaxing targeted non-anticipativity constraints, while maintaining constraints in scenario group sub-problems. Vayanos et al., (2011) use approximations to solve the otherwise computationally intractable problems of endogenous uncertainty as the problems increase in size. Their method works for continuous distributions of the uncertainty space.

Mercier et al., (2008) further refine the definition of particular subset of decision-dependent uncertainty planning problems, in which the *observation* of an uncertainty is dependent upon a decision (e.g. to invest in a research project), but the *outcome* of an uncertainty is exogenously defined. The pharmaceutical research and development problem of Colvin et al., (2008, 2010) is a good example of this. They develop a multi-stage stochastic programming formulation for research and development projects of consecutive stages. The model computes the optimal strategy which maximises the expected net present value of the research portfolio. The model incorporates revenues, project costs, project durations and uncertainty in the success of the research project stages. The assumption of consecutive research project stages allows a reduction in the size of the **scenario tree**. These papers incorporate decision-dependent uncertainty, where the

outcome of a stage of the research project is only observed at the completion of the stage. The probability of a research project completing successfully is the product of the probabilities of the individual stages. These probabilities are defined exogenously. Stages of separate projects can be conducted in parallel, subject to resource constraints, but stages of the same research project must be conducted consecutively. The authors develop a branch-and-bound algorithm which only impose [NACs](#) when they are violated within the search tree, thus reducing the size of the initial problem. Solak et al., (2010) relaxes some of the assumptions in the work of Colvin et al., (2008, 2010), incorporating stochastic R&D investment requirements and continuous investment decisions.

2.3.2 *Stochastic Dynamic Programming*

[Stochastic dynamic programming \(SDP\)](#) is a general mathematical modelling framework that can be used to model sequential decisions under uncertainty. However, in its original form, as outlined by Bellman, (1952), [SDP](#) is capable only of modelling problems of a limited size. This is due to the three curses of dimensionality in the state, action and information spaces (Powell, 2011), that very quickly result in computationally intractable problems as the numbers of uncertain parameters and decisions increase.

A dynamic program is formulated as a sequence of decisions made in each time-step. The dynamic program is then solved to find an optimal sequence of decisions, or policy, that result in a maximum (or minimum) objective function. The objective function in each time period is the sum of the cost of being in the pre-decision state, the cost of making a decision, conditional on the pre-decision state, and the expected future value of being in the post-decision state. In each time period, new information may arrive which affects the pre-decision state and the actions in the subsequent time period.

In the dynamic programming version of the decision-dependent uncertainty problem the state space explodes in size for the same fundamental reason as stochastic programming, creating a problem too large to be solved analytically, or numerically exactly. [ADP](#) (reinforcement learning) shows promise for tackling such problems (Bertsekas, 2005). Research in this area is growing but the techniques are host to computational issues that are problem specific. The main advantage of the [Approximate Dynamic Programming \(ADP\)](#) approach is that continuous decisions and uncertainties, and detailed models of complex phenomena can easily be modelled. The difficulty is in designing algorithms that converge to an optimal solution.

Webster et al., (2012) investigate decision-dependent technology uncertainty in a multi-stage sequential decision process approximating R&D. They apply [ADP](#) techniques to solve a modified version of the

IAM DICE formulated as a stochastic dynamic programme. The conclusion is that models which incorporate an exogenous treatment of technical uncertainty, i.e. ignoring the decision dependent effects of technological change, are likely to underestimate the effect of uncertainty on the decisions.

2.3.3 *Stochastic Decision Trees*

In simple cases (i.e. simpler than for an **ESOM**), the decisions of a dynamic program can be formulated as a decision tree, each branch of which is folded back to its node from the end to beginning, with the 'best' decision chosen in each branch. This approach is only tractable in the simplest of cases, when the number of time periods is small (for example in a 2-stage problem). A decision tree can then be used on its own to manually calculate optimal decisions under uncertainty.

Real options is an alternative way of formulating investment decisions to that of simple cost-benefit analysis, which draws upon the theory of financial options. A real option exists when a decision is irreversible, and there is the ability to delay an investment as an alternative to investing today.

In chapter 5 of Dixit et al., (1994), a real-options example is given in which the decision to invest in a nuclear power station involves both endogenous and exogenous uncertainty (labelled 'technical' and 'input' uncertainty respectively). The example uses a continuous decision and uncertainty space, but by using discrete distributions, a simplified version of the same problem can be represented using a stochastic decision tree.

Continuing with the real-options examples, Eckhause et al., (2009) and Eckhause et al., (2014) demonstrate, for the selection of CCS R&D projects, the value added from a real-options approach over a now-or-never policy. The decision to invest in one stage of a multistage R&D project is equivalent to the purchase of a call option. The option is exercised upon continuation to the subsequent stage (i.e. upon observing successful completion of the previous stage), or abandoned if the previous stage fails.

2.3.4 *Selecting a Modelling Approach for the Learning Model*

Above, I discussed how **dynamic uncertainty** has been treated using two main approaches: stochastic programming and (stochastic) dynamic programming. While both of these approaches look promising, the former emphasises ease of solution with some concessions to realism, while the latter allows very detailed modelling of processes, but requires bespoke solution algorithms to solve the models. The approach taken in this thesis is the former, stochastic programming, to leverage the commercial solvers available and to concentrate upon

the dynamics of the problem, rather than the development of a novel solution algorithm. Within stochastic programming, there are still a number of different approaches that are possible, as I touched on in Section 2.3.1. The approaches differ in the type of uncertainty that is modelled. Gupta et al., (2014) distinguish between two types of **endogenous uncertainty**, where in Type I, decisions affect either the probability distribution, or in Type II the timing of the resolution of the uncertainty is decision dependent. In this thesis, I focus on Type II endogenous uncertainty, where the observation of the outcome of uncertain research projects is dependent upon whether investment takes place in the research project.

2.4 LEARNING: REDUCING UNCERTAINTY, INCREASING KNOWLEDGE

In this thesis, **learning** refers specifically to **epistemic uncertainty** about parameters in an **ESOM**, such as technology cost, and is modelled in an endogenous and sequential manner — the true value of an exogenously defined uncertainty is observed conditional upon a decision to invest in a research project. The observed outcome (resolution) of the uncertainty influences the subsequent recourse action. Incorporating learning into a model means that either the timing of the observation of an uncertainty (Type II endogenous uncertainty), or the ‘shape’ of uncertainty (Type I endogenous uncertainty) is itself uncertain, or the function of a decision. In the first case a double hedge is required in the presence of irreversible effects, or the uncertainty must be modelled as decision-dependent.

Learning, as a concept, has been studied in detail since the 60s (for example, by Arrow, 1962; Hartley, 1965). Learning in this context was initially used to explain the increasing economic gains associated with the accumulation of knowledge¹; otherwise known as learning-by-doing. The recent **learning** literature describes the influence of both rates and direction of learning in models of dynamic climate-economic systems. Oppenheimer et al., (2008) propose the term “**negative learning**” when learning takes place leading to a false scientific consensus away from the true value of an important parameter, such as climate sensitivity, and examine the cost of the resultant (imperfect) policy. Kriegler, (2009) expands upon this definition to discriminate between two types of learning under **ignorance**. Type A, in which the set of known models is expanded, and type B, in which

¹ Note that knowledge accumulation is the counterpart of **epistemic uncertainty**, which is uncertainty due to incomplete knowledge. If learning is taken to be an increase in knowledge, it holds that learning corresponds to a reduction in **epistemic uncertainty**. If probability distributions are used to represent an **epistemic uncertainty**, then a definition of ‘reducing uncertainty’ is required, because an investment in learning may not result in a reduction in uncertainty if it is measured using a normalised measure, such as (90th-10th)/50th percentiles (Nemet et al., 2016)

we update our beliefs about a known model. While type B learning may be represented through Bayesian updating, type A may only be informed through Bayesian inference since type A learning refers to the "...emergence of a positive belief in an area of the model space that was not supported by the prior belief." Lange et al., (2008) show that uncertainty and learning exert opposing forces upon welfare in a standard two-stage expected utility model.

Lorenz et al., (2011) adapt the [IAM MIND](#) to a two-stage decision problem formulated as a recursive optimisation problem. They distinguish between learning and anticipation of learning, resulting in a similar hedging effect to a two-stage stochastic optimisation problem in which emission trajectories differ according to the anticipation of future learning of climate sensitivity and damage amplitude. They show that the anticipation of learning in the future can be evaluated, just as the value of actually learning can be. Hannart et al., (2013) refine the definition of learning, through defining the complement of negative learning - disconcerting learning. This is when learning reduces bias (progressive learning), but when overall uncertainty increases (disconcerting progressive learning). Hannart et al., (2013) also observe that Keller et al., (2004) and Webster et al., (2008) both investigate the difference between decision making under static uncertainty, and decision making under [dynamic uncertainty](#) (where learning takes place). Both studies conclude that optimal decisions adapt in the presence of an irreversible effect, but take no different action under smoothly increasing costs, such as damage costs. This makes intuitive sense, as a decision model will opt to hedge against future uncertainties, if the hedge is cheaper than the cost of the .

Kolstad, (1996) defines three mechanisms for learning:

- experimental - such as perturbing emissions to see the effect upon climate
- purchased - such as investing in R&D to increase knowledge
- passive - waiting to see what happens

A fourth strategy is to ignore the ability to learn and just act now.

Powell et al., (2012) describes two categories of learning; off-line and on-line. Off-line learning occurs before a decision is made. For example, an epistemic uncertainty is identified, a research project conducted to (perhaps partially) resolve the uncertainty, and subsequently a decision is made based on the new information. On-line learning occurs in parallel with the decision making processes, and can interact with the decisions being made. For example, a decision maker could continuously monitor the results of the decisions. Each separate decision, such as setting the price of an airline seat, results in a new piece of information, such as a sales figures. This information is used to influence the airline price in real-time. Off-line learning

is more aligned with the long-term decision processes in this thesis, while on-line learning is more suited to algorithmic approaches to price setting, machine learning and optimal control.

This discussion of learning should not be confused with [ETL](#), the techno-economic theory of cost reduction in technologies due to experience².

2.4.1 *Learning and R&D in the energy sector*

Previous studies of the energy system have explored some aspects of learning. [Stochastic programming](#) explores the trade-off between acting now or waiting until an uncertainty is resolved. Decision theoretic approaches explore the option to invest in information about an uncertainty (a prospecting problem).

Eckhause et al., (2009) and Eckhause et al., (2014) model CCS R&D projects using a real-options approach. The probabilities of a project completing successfully are known prior to the computation of the model, and so the formulation is functionally equivalent to the earlier definition of Mercier et al., (2008) in which the observation of an uncertainty is dependent upon a decision. However, the objective function differs to that of a stochastic programming formulation, as the aim is to maximise the probability of at least one successful R&D project subject to budget constraints.

Baker et al., (2006) investigate the difference between R&D programmes which target minor improvements in existing technologies (such as incremental efficiency gains in a coal fired power station) versus dramatic cost reductions in novel technologies (such as solar photovoltaics). The analysis reveals the effects of R&D to be ambiguous in terms of whether the projects act as a hedge against other sources of uncertainty. This work is extended by Baker et al., (2008) who again use the [DICE](#) model to explore the effect of R&D funding upon the marginal abatement cost curve. The authors distinguish between incremental and risky R&D projects, the latter associated with technological breakthroughs, but also high probabilities of failure and high cost, whereas the former reflects incremental improvements in existing (fossil fuel) technologies. Their results show that the type of R&D project depends upon the probability distribution used for climate damages. In other words, the amount of carbon that is able to be emitted in the future either closes off options for conventional fossil fuel plant, even very efficient technologies, under high-damage, or under a low-damage case, disincentivises investment into breakthrough zero-carbon technologies.

Baker et al., (2011b) look at uncertain technology learning (as in [ETL](#)). While this work does not use an [ESOM](#), rather a stylised por-

² [ETL](#) assumes a fixed learning-coefficient which represents an extrapolated relationship between installed capacity of a technology and its cost.

trayal of technology value through marginal abatement cost curves derived from MiniCAM, an IAM. Importantly, these MAC curves do take into account interactions. However, the paper does not model decision-dependent uncertainty, and treats the uncertain outcomes from research as an *exogenous uncertainty* modelled using a two-stage stochastic programme.

Baker et al., (2015) link expert elicitations performed over a five year period on the relationship between investment in R&D and technological learning, with optimal decision making, through an IAM called GCAM³. The GCAM model is used to estimate the economic payoff of a particular outcome of technological change. In their framework, Baker et al. assume that technological outcomes of three levels of R&D investment are uncertain, with probability distributions representing outcomes for eight technology attributes (including cost and/or efficiency for solar, nuclear, CCS, biomass and bio-electricity). An interesting insight from the paper concerns the value of technologies which are flexible under different out-turns. Specifically, identifying technologies which avoid becoming stranded assets under worse-than-expected outcomes.

Baker et al., (2011c) show how R&D has a value not just in lowering costs of a technology, but reducing the uncertainty surrounding whether the costs can be lowered, and by how much. In other words, the expected value of better information - that is knowing the true distribution of uncertainty instead of estimating this uncertainty - is very large, especially for technologies with very large research budgets, such as nuclear power.

Nordhaus, (2014) argues that learning models of technological change which rely upon assumed learning curves, or learning coefficients, are fundamentally biased. This is because it is difficult to distinguish between exogenous technical change and endogenous learning effects, the learning parameters are not robust to alternative specifications, and that overestimates of learning coefficients in optimisation models will lead to an underestimate of the marginal cost of output leading to massive investment.

A new paper, Santen et al., (2016), offers a framework which directly addresses the concerns of Nordhaus, (2014). Santen et al., (2016) model the uncertainty surrounding endogenous technological change, as the induced process of the uncertain outcomes from decision-dependent RD&D investments. This avoids the need for unknowable (and sensitive) learning parameters, by modelling the mechanism by which RD&D influences the cost of technologies using Type I endogenous

³ Note that GCAM is the subject of a sensitivity analyses performed in McJeon et al., (2011) and Scott et al., (2014) on page 26

uncertainty⁴. Also, this approach exhibits an understandable uncertainty structure, for which it is possible to elicit beliefs from experts.

2.4.2 Value of Information

Studies of the **VOI** rank uncertainties according to the benefit within the model of: removing the uncertainty (**EVPI**), or ignoring uncertainty (**EVIU**). When considering uncertainty within decision support frameworks, quantifying **VOI** is a logical extension to the quantification of uncertainty. For example, Felli et al., (1999) perform a Bayesian uncertainty analysis of a health model, where model outputs correspond to a probability weighted cost function (expected cost). Any change in input distributions through a learning process can be quantified in the change in expected cost. The difference in the expected cost when perfect information is available (which may be aspirational) versus imperfect information, gives the expected value of perfect information (**EVPI**). And Popp et al., (1997) estimate the value of information associated with key climate parameters using the **PRICE** model. They quantify the value of information as between \$1 to \$2 billion for each year the resolution of the uncertainty is moved towards the present. Studies that have used static representations of uncertainty in two-stage stochastic programming models have used **EVPI** to quantify the value of removing the uncertainty from the model (Krukanont et al., 2007).

Morgan et al., (1992, ch.12) compares **EVPI** with **EVIU**. Both compare expected value of a Bayes decision with another decision made without uncertainty. For **EVIU**, the other decision is made when uncertainty is ignored. For **EVPI**, the other decision is made when uncertainty is removed, by obtaining perfect information. Total **EVPI** is the **EVPI** when all uncertainties are included. Partial **EVPI** is when only a subset of uncertainties are computed. **EVPI** extends (exogenous) **sensitivity analysis** to evaluate the value (or utility) to a decision maker of reducing uncertainty in a model input, rather than merely identifying those inputs to a model to which the outputs are sensitive. For example, Brennan et al., (2007) computes partial **EVPI** to generate a rank of a subset of uncertain parameters in a health-economic decision model. Conejo et al., (2010) describes the calculation of **EVPI** and **Value of Stochastic Solution (VSS)** when using **stochastic programming** techniques. In a two-stage stochastic programme, **EVPI** is computed as the difference between the probability weighted stochastic solution and the deterministic solutions of each branch of the **scenario tree**. Within a multi-stage stochastic programme, the computation of **EVPI** is more complicated; the **scenario tree** is more complex due to

⁴ Increasing investment in RD&D increases the likelihood of success (i.e. changes the probability distribution) of solar PV capital cost. See Table 12 on page 97 for the distinction between Type I and Type II endogenous uncertainty

the multiple recourse stages. Alternatively, Baker et al., (2011c) define **EVBI**. Modelling methods which include dynamic uncertainties incorporate the value of information into the decision framework. These frameworks are referred to as online- or offline-learning or decision dependent models.

2.5 BRINGING ALL THESE ASPECTS TOGETHER

In this chapter, I reviewed the literature underpinning the topics covered in this thesis. I now discuss the links between the chapters given the context laid out in this review of the literature.

2.5.1 *The Links Between Chapters*

In Section 2.1, I showed that **ESOMs** have not yet been used to their full potential to explore **parametric uncertainty** and **structural uncertainty**. Aside from the very first **ESOM** papers in the 70s and 80s, which were concerned with the new methodological insights from the approach, the role of uncertainty has received a consistent level of attention over the decades. With a few exceptions, this attention has not resulted in the use of systematic computational techniques, such as global sensitivity analysis, to understand the influence of uncertainty on the model results. The focus instead has been on the definition and comparison of an increasing quantity of scenarios. Where uncertainty has been acknowledged, **parametric uncertainties** have been the predominant focus. The notable effort of the model comparison exercises, such as those of **EMF**, have helped generate an understanding of the structural differences between models, but there is little evidence of a comprehensive understanding of the influence of **structural uncertainty** within this field. As computational power has increased, some studies have made use of a very large number of scenarios, but often these are not conducted in a systematic manner, rather guided heuristically through interactions between modeller and client (such as in Usher et al., 2010a, for example). Where **parametric uncertainties** were investigated, invariably the studies focused on technology costs, emission constraints and resource prices. But other aspects of the energy system are also uncertain. Few studies have investigated the importance of non-technology uncertainties in a way that relates easily to the input parameters of a model.

The focus on scenario techniques in **ESOMs** is quite correct, given that the nature of the uncertainties for many parameters requires an approach which allows for the imprecise knowledge regarding probabilities (using the terminology of Stirling, 2007). The nature of these uncertainties need to be taken into account when selecting a global sensitivity analysis technique, as stated in the first research question.

Little evidence was found of the use of the global sensitivity analysis techniques, outlined in Section 2.2. However, some of the [sensitivity analysis](#) techniques have been used in the related field of integrated assessment modelling, showing that these techniques do have potential within this field as the model formulations of the two fields are often similar. In Chapter 3, I investigate and select an appropriate sensitivity analysis technique, the Method of Morris, for use with energy system models, presenting the results of a global sensitivity analysis of the [ETI-ESME](#) model in the following Chapter 4. The technique chosen could be applied to other [ESOMs](#), which could then give greater insights into the structural differences between the models. Understanding that the results from the analysis are a function of the [ETI-ESME](#) model, the [global sensitivity analysis](#) identifies some critically uncertain parameters to the UK energy system. A subset of these are dynamic uncertainties and are suitable for implementing into a learning model.

The core concepts of [dynamic uncertainty](#): learning about critical uncertainties at a national scale, the intersection of these uncertainties and [agency](#), and the ability to learn about the uncertainties, are tackled in Chapter 5. Several studies indicate that the act of deciding when and how much to invest in learning is important, and that this importance is quantifiable. In the learning literature, the focus is to investigate and quantify the value of reducing epistemic uncertainty, or the economic benefit (or cost) of scientific progress. In Chapter 5 I apply an existing multi-stage mixed-integer stochastic programming formulation for learning about critical energy system uncertainties in a novel setting - linking investment in a portfolio of research projects to the critical uncertainties identified by the [global sensitivity analysis](#) of the [ETI-ESME](#) model. The results from this approach are presented in Chapter 6.

One strength of energy system models is to highlight clusters of technologies across the whole energy system which work together. This attribute of a system-approach to energy planning allows modelers to identify key activating system conditions. However, these system conditions are often sensitive to numerous uncertainties, directly and indirectly related to the technologies themselves, many of which could be resolved through investment, experimentation or research. Coupling the system-wide perspective of an [ESOM](#) with a treatment of [dynamic uncertainty](#) offers new insights into cost-effective decarbonisation of national energy systems. Thus the results presented in Chapter 6 link the optimal research strategies for key uncertainties in the energy system with the optimal composition of the energy system given the success or failure to resolve these uncertainties (see Section 6.4). The additional value of resolving these uncertainties is explored through the maximum expected cost-threshold a rational investor would be willing to pay.

SENSITIVITY ANALYSIS OF AN ENERGY SYSTEM MODEL

Section 2.1 introduced the various methods that have been used to manage uncertainty in [Energy System Optimisation Models \(ESOMs\)](#). Section 2.2 presented a review of the available sensitivity analysis techniques available. Here, I bring these two branches of the literature together, and develop a sensitivity analysis methodology for an [ESOM](#).

As a multidisciplinary domain involving environmental, engineering, and economic elements, the energy sector has a long history of the application of a broad range of large-scale complex models (Jebaraj et al., 2006). A key subset of these models are [ESOMs](#), which are a class of linear optimisation models. These models minimise selected (discounted) cost metrics under resource, technological, social, and policy constraints. [ESOMs](#) are frequently used to support national government policy making (Strachan et al., 2008b), and at the global level, the International Energy Agency (IEA) use [ESOMs](#) to inform their headline energy strategy publications (International Energy Agency, 2012).

In the UK, [ESME](#), a UK focused energy system model developed by the [ETI](#), has been used for identifying opportunities for the funding of research. In the first half of this chapter, I introduce the particulars of the [ESME](#) model. This is followed by a typology of the model inputs and a description of the model data requirements. In the context of [ESOMs](#), Section 3.2 presents a more technical discussion of sensitivity analysis than Section 2.2 and outlines the advantages and disadvantages of the two main branches of methods: local and global approaches. While local approaches are quick and easy to perform, the results can be misleading particularly for models that are non-linear and non-additive (see page 42 for a discussion of the linearity of [ESOMs](#)). In contrast, global approaches are computationally demanding, but offer an array of robust indices for the ranking and comparison of inputs. Given the unique combination of elements that make up the discipline of energy system modelling, variance-based sensitivity approaches which require the definition of probability distributions and are computationally intensive are not appropriate. The technique of [scenario analysis](#), used in energy system modelling to manage the [Knightian uncertainty](#) applicable to many of the inputs used, also precludes the use of probability distributions which assume of risk based approach.

Section 3.3 outlines how the Method of Morris, a difference-based global approach which computes elementary effects — a ratio of the change in output for each input based upon random permutations of all other inputs — is deemed particularly appropriate for conducting a [global sensitivity analysis](#) of ESME. Section 3.3 goes on to describe this method, and introduces an novel extension to this method to enable better sampling of a model’s input space, discusses the suitability of the technique and various implications of the method. The technical elements of this section are supported by a nomenclature, available on page xxv, and the formulation developed by Morris, (1991) and extended by Campolongo et al., (2007) in Appendix A.

Section 3.4 then discusses the challenges in linking the outputs from the sensitivity analysis to the data requirements of the learning model introduced in Chapter 5. In particular, harmonising the dynamic aspects of the learning model with the non-temporally disaggregated inputs of the sensitivity analysis make extracting direct insights from the sensitivity analysis challenging. Section 3.4 also introduces a key framework for assigning links between the nature of an uncertain input and the suitability of an input for use in the learning model introduced in Chapter 5. This framework discerns whether the uncertainty associated with an input parameter may be reduced, and in what way, and if not, what methodology is appropriate. In Section 3.4.4, I conclude with a discussion of some of the methodological issues associated with running a sensitivity analysis using the Method of Morris on energy optimisation models.

3.1 THE ETI-ESME MODEL

[Energy System Modelling Environment \(ESME\)](#)¹ was developed to provide a decision support solution to the Energy Technologies Institute, a private-public partnership between the UK Government, and a consortium of large energy utilities and industry including EDF-Energy, Shell, BP, E.On, Rolls-Royce and Caterpillar. The objective of the ESME model is to identify key technologies for the UK energy system out to 2050. The focus is particularly upon technologies that could be targeted through research and development funding, as ETI is a funder of research. Thus the outputs of the model are used to inform the direction of funds that ETI controls to assist in achieving the transition to a low-carbon energy system. In addition, the ESME model is used to provide an evidence base to Government and research bodies as to where funding could be most effective. ESME is a modelling tool designed by the ETI, populated with data that represents the current UK energy system, together with a characterisation of future technologies and projections of demands for energy services.

¹ http://www.eti.co.uk/technology_strategy/energy_systems_modelling_environment/

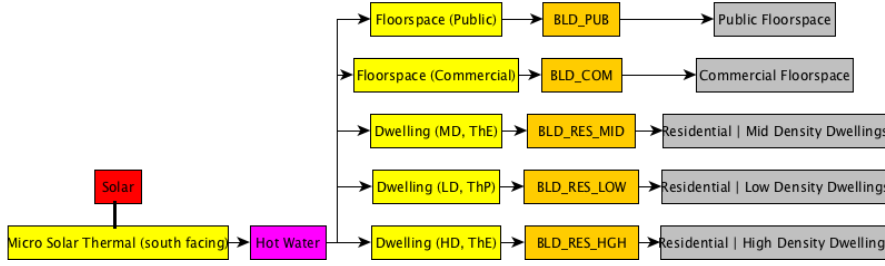


Figure 1: A diagram of a small portion of the ESME model. The solar resource (red) is consumed in a solar collector conversion technology (yellow) to generate a hot water resource (purple). The hot water is converted in Dwelling conversion technologies (yellow) to produce the hot water energy service (orange) to meet the energy service demands in dwellings of varying densities (grey)

3.1.1 The Objective Function

As an optimisation model, the key output variable of interest is that of ‘total energy system cost’ — this is the objective function that is minimised by solving a system of linear equations using a commercial solver software package. In the case of *ESME*, the objective function is the discounted sum over all time periods of the cost of investing in and operating all resources, plants, infrastructures and demand technologies in the UK energy system, as shown in Equations 1 to 5.

$$\min \sum_t^T \left(\sum_k^K C_{t,k}^{Cap} + C_{t,k}^{Fix} + C_{t,k}^{VOM} \right) \quad (1)$$

$$+ \left(\sum_l^L C_{t,l}^{Cap} \right) \quad (2)$$

$$+ \left(\sum_g^G C_{t,g}^{Cap} + C_{t,g}^{Fix} + C_{t,g}^{VOM} \right) \quad (3)$$

$$+ \left(\sum_{x \in Tra}^X C_{t,x}^{Cap} + C_{t,x}^{Fix} + C_{t,x}^{Flow} \right) \quad (4)$$

$$+ \left(\sum_{x \in Res}^X C_{t,x}^{Flow} \right) \quad (5)$$

where

$C_{t,a}^{Cap}$ Capital cost at time period t , for index $a \in k, l, g, x$

$C_{t,a}^{Fix}$ Fixed cost at time period t , for index $a \in k, l, g, x$

$C_{t,a}^{VOM}$ Variable operation & maintenance costs at time period t , for index $a \in k, l, g, x$

$C_{t,a}^{Flow}$ Flow cost at time period t , for index $a \in k, l, g, x$

- t Set of time periods in 5-year steps from 2000 to 2050
- k Set of conversion technologies
- l Set of retrofit technologies
- g Set of storage technologies
- x Set of products, where the subset $Tr a$ includes transmittable products (such as electricity and natural gas) and the subset Res include all resources that can be purchased as imports (such as petrol and diesel, coal etc.)

There are two alternative formulations of the objective function. In this thesis, the ‘standard’ version of the model is used for the majority of the work, although the ‘elastic demand’ version is used to compare with the findings from another study (see Section 4.3).

In the standard version of the model, [ESME](#)’s objective function is to minimise the discounted total energy system cost. Demands for energy services are inelastic and so must be supplied at any cost. The model can choose from a database of technologies, each of which consumes different resources. The model results include the least cost energy system and the portfolio of technologies which comprise that energy system. Imposing constraints, such as environmental (e.g. an emissions cap), or growth constraints, (e.g. the rate at which new technologies can penetrate the technology portfolio), further shape the results.

In the elastic demand version of the model, the objective function is modified to compute a proxy for welfare, by maximising the consumer/producer surplus in a policy case versus a reference case. Own-price elasticities are defined for each energy service demand. The values used are derived from a literature search outlined in [Pye et al., 2014](#). Cross-price elasticities are assumed to be zero. Demand curves are extrapolated from the shadow prices of the energy balance constraint in a reference case run of the model (i.e. in the absence of an emission constraint). Then, in a policy case, the model may choose to reduce the demand for energy services, incurring losses in consumer welfare, if that is cheaper than increasing the size of the supply side. The change in consumer/producer welfare between the reference and policy cases captures more of the economic effects of energy. This provides useful insights into the role that demand side management could play in future decarbonisation scenarios. However, own-price elasticities for energy services are themselves uncertain ([Pye et al., 2014](#)).

The [ESME](#) model is formulated as a linear programme. The term ‘linear’ refers to the fact that decision variables are not multiplied with one another, thus the objective function is linear. The objective function is subject to linear inequalities and linear inequality constraints. A common misconception is that a linear program can only

model linear phenomena. But this is not the case, as piecewise linear functions can be used to approximate non-linear cost curves (for example), and bundles of constraints often result in an implicitly non-linear supply-cost curve (such as max build rates and quantities). However, these individual functions must be monotonic and convex/-concave.

3.1.2 Input Data

The inputs to the model are treated in a number of different ways according to the nature of the uncertainty associated with the parameters. At the highest level, are demand scenarios. Technologies and other parameters may be specified using probability distributions.

An important abstraction in energy system models is to use the demand for energy services rather than demand for final energy. This enables the model to choose demand technologies as well as supply technologies. For example, more efficient refrigerators can be chosen, or even a large-scale district heating system to replace individual house boilers. Either of the latter technologies may supply the energy service ‘domestic heating’.

Quantifying and projecting forward energy service demands is done through modelling itself. The magnitude of energy service demands (or end use service demands to use the ESME terminology) are very important drivers of the size of the energy system, particularly in the default inelastic demand mode. Energy service demands are varied as part of the package of assumptions contained within a scenario. In ESME, the values are not able to be varied probabilistically.

Input	Description	Example	Indices
Prices	Price of resources	Diesel Price	x, b, i
Costs	Cost of technology	Capital cost of nuclear power station	$k/g, b, i$
Demand	Amount of energy service demanded	Medium Density Residential Hot Water	x, b, i
Constraints	Technology Build Rates	Maximum Rate at which nuclear power can be built	$k/g, b, i$

Table 3: List of inputs to the ESME model

There are 29 different energy service demands, divided among transport, residential, commercial/public sector and industry. The demands in the industry sector are simplified, largely because the industrial sector is so heterogeneous. This sector is subdivided into 8 sub-sectors, each of which deals with a particular category of energy demand, such as high-temperature heat, motive force, etc. The residential sector is disaggregated into three dwelling densities to re-

Input Category	Entries	Parameters	Uncertain
Conversion Technologies	75	6	1
Transmission Technologies	19	8	1
Industry	32	4	0
Building	23	6	0
Transport	50	9	1
Resource	10	2	2
Constraints	≈ 50	1	0 to 1

Table 4: Breakdown of the categories of inputs in the [ESME](#) model, the number of parameters in each category, and the subset of parameters available for uncertainty quantification. The total available uncertain parameters for sensitivity analysis numbered around 200.

flect the different patterns of energy use across urban, peri-urban and rural housing. Energy service demands for appliances, cooking and air conditioning are separate. Transport is disaggregated into aviation, rail, road and maritime. Road is the largest service demand and is further divided among vehicle types, such as passenger car, heavy, medium and light goods vehicles. A selection of sample inputs to [ESME](#) are listed in Table 3. A summary table of the total number of inputs to the [ESME](#) are listed in Table 4.

3.1.3 High-level Scenarios

The scenarios contained in the ETI-ESME model are based on a high-level qualitative narrative constructed by [ETI](#)'s stakeholders. Typical scenarios concern the drivers of energy service demands, such as changes in the economic and demographic makeup of the UK. Resource prices also change across scenarios, representing differing patterns of geopolitical and economic changes that are well outside the boundaries of the model.

3.1.4 Monte Carlo Feature

One of the attractive features of the [ESME](#) model is that it has been designed to operate using Monte Carlo sampling, as mentioned in Section 2.1.3. This makes it very easy to run the model over a number of different parameter values, potentially allowing a thorough exploration of the parameter space. Conveniently, this is also advantageous for running a sensitivity analysis.

3.1.5 *Key Decision Variables*

The key [decision variables](#) are firstly the technology capacity to install in each time period, and secondly, the production of the available capacity within each time-slice of that period. The model combines capacity expansion with dispatch of plant across the energy system.

3.1.6 *Spatial and Temporal Aggregation*

The degree to which the model represents the energy system in spatial and temporal disaggregation is a major design decision. Increasing the model resolution can dramatically increase solution time and data requirements, with decreasing returns to the time and effort required. The ESME model has twenty-four regions. Three are for storage technologies only, 19 are onshore demand and supply nodes, and the remainder are offshore supply nodes, mainly for offshore renewable technologies. The location of demand and supply has implications for the design of a transmission system, and other infrastructure. ESME does a reasonable job of representing the spatial key aspects of the UK energy system, and is better than competing UK energy system models at this role, given its system overview. The ESME model has 10 time slices. The year is divided into two seasons, and the peak of the year is explicitly represented. Within the two seasons and the peak, there are five periods corresponding to various hours in the day to adequately represent the load curve across the population. Each supply and demand technology has its availability or load factor disaggregated across these or less disaggregated time slices.

3.1.7 *Elastic Demand Formulation*

Elastic demands enable the ESME model to choose the optimal reduction in energy service demands as an alternative to investing in supply side capacity. This can be an important consideration under decarbonisation scenarios, as the cost of mitigating emissions from the supply side may be considerably more expensive than reducing demand. In an economic sense, the welfare cost of reducing demand maybe smaller than the welfare cost of increasing investment in low-carbon technologies to meet fixed (inelastic) demand for a given energy service.

During the course of the PhD research, I helped develop an elastic demand component to the ETI-ESME model, which was subsequently used in Pye et al., (2014). The original formulation of the [ESME](#) model treats demand for energy services as inelastic, thus supply of energy must always meet demand for the downstream energy services. Under the elastic demand formulation, demand for energy services can change endogenously according to the change in price of the energy

services. The elastic demands are formulated as a piecewise linear approximation of the demand curve. While the elasticities are defined exogenously, the base price from which the demand curve is extrapolated using the elasticity, must first be generated from a reference scenario. The objective function is modified to include the consumer-/producer surplus computed as a result of the change in demands for energy services. Thus the model includes a measure of the change in welfare instead of just the cost of the energy system.

Pye et al., (2014) reviews the literature for own-price elasticities and performs an uncertainty analysis upon the elasticity values. The key findings are that the majority of demand reduction occurs in the transport sector, and that over the range of values taken by the elasticities, demand reduction is a consistently important mitigation option under a carbon constraint.

3.2 SELECTION OF AN APPROPRIATE SENSITIVITY ANALYSIS APPROACH

In this section, I determine which sensitivity analysis method is appropriate for analysing the ESME model. In contrast to Section 2.2 of the previous chapter, which provided an overview of the sensitivity techniques available, the focus here is upon determining which approach is most suited to the characteristics of the ESME model.

Table 5 gives a summary of the characteristics of five different methods of sensitivity analysis. The variance based approach, at over 10 days of computation time, is bordering infeasible, while the others trade off the length of computation against the insights gained.

3.2.1 *Matching a sensitivity analysis approach to an energy system model*

The work of Saltelli et al., (2008a) outlines best-practice for a sensitivity analysis of a model. In the setting of an energy system model, where the uncertainty surrounding the majority of inputs is managed using *scenario analysis*, an uncertainty analysis is not applicable because there is no likelihood associated with the values of any one particular input. As such I focus on determining the most appropriate methodology for a sensitivity analysis given the known characteristics of energy system models. This subsection tackles the following issues in turn:

- ESME has 200 uncertain input parameters
- the data associated with many of the uncertain inputs is not probabilistic, although plausible ranges are feasible for all inputs,
- the model is computationally demanding with run-times measured in minutes

Table 5: Characteristics of sensitivity and uncertainty analysis methods (adapted from Flechsig et al., 2012; Saltelli et al., 2008a) given an upper bound on CPU time (3 days)

Type	Morris	Variance	Factorial	MCF ³	OAT
Model independent?	yes	yes	yes	yes	yes
Sample source	levels	dist's	levels	dist's	levels
No. factors ²	20 – 100 ¹	< 5 ¹	> 1000 ¹	< 5	< 500
Factor range	global	global	global	global	local
Multi-factor variation	yes	yes	yes	yes	no
Correlated factors?	no	no	yes	yes	no
Cost (for k factors)?	10(k + 1)	500(k + 2)	k → 2k	500 + 1	2(k + 1)

¹ using groups of factors would enable larger numbers of factors to be explored

² values represent an approximate maximum in the given time threshold, but not a hard upper limit for the approach. Assumes 5 minutes per simulation and 30 groups of factors and no parallel computation

³ MCF=Monte Carlo filtering

- the model results can be non-linear in response to changes in inputs

The nature of the [ETI-ESME](#) model is that it is a large, complex model, with around 200 input data points predominantly based on projections or assumptions rather than empirical data. The setting in which we wish to conduct the sensitivity analysis is to identify the factors (input parameters) to which the model is most sensitive (factor prioritisation), as this implies that reducing the uncertainty surrounding these sensitive parameters, perhaps by investing in a research project, would be particularly valuable. It would also be useful to know which input uncertainties that are not amenable to learning are important for the model outputs. Finally, for the purposes of reducing the number of parameters to include in a [scenario analysis](#), it is useful to identify those parameters that do not affect the output of interest and can thus be safely fixed to a value within their range. The large number of inputs suggest an approach which is computationally cheap, or allows the grouping of parameters to reduce the number of simulations required.

There are 200 input parameters of interest (see [Table 4](#)). These include conversion technologies (of which 75 are included in the [ESME](#) model) related parameters such as capital costs of generation technologies, technology efficiencies, lifetimes, build-rate constraints², and [hurdle rates](#). Resource related parameters include domestic resource production costs, or the traded market price for resource imports, and constraints upon the availability of domestic resources.

Similarly, there are a large number of outputs from the model. As an optimisation model, the objective function which is minimised is the total system cost. Other outputs include the size of the electricity system (power capacity and production by technology), CO₂ price and reduction in energy service demands in response to changes in price. A more detailed discussion of output metrics which align with the data required for the learning model can be found in [Section 3.4.3](#).

As in other [ESOMs](#) used under a scenario approach, no measure of likelihood is associated with the values of many inputs to the [ESME](#) model. The nature of the uncertainty associated with inputs that represent distant and unquantifiable phenomena mean that such unknowns are not amenable to the use of probability distributions. However, in most cases, plausible ranges of values, and upper and lower bound that is deemed reasonable, can be used in line with Laplace's criterion of insufficient reason (we know absolutely nothing) and can therefore assume a uniform distribution. More discussion on this assumption is included in [Section 3.3.3](#).

Another challenge relates to the likely correlation of input parameters. Only factorial and Monte Carlo filtering sensitivity analysis meth-

² In [ESME](#), build rate constraints are modelled as the upper bound on new capacity that can be installed per technology or group of technologies in each year

ods allow for the representation of a joint input distribution, while the Method of Morris, Variable Based and one-at-a-time approaches make the potentially strong assumption of independence of inputs. In the context of a sensitivity analysis, what is the importance of capturing input correlations? A primary aim of the analysis is establish some insight into the ‘behaviour’ of the model. This behaviour maybe a function of both input parameter values and the formulation. Assuming independence between inputs reveals any interaction effects in the outputs as solely due to the model formulation. This may be a way of improving the transparency of a model. The risk of ignoring the correlation between inputs is that the sensitivity analysis wrongly identifies an influential input as non-influential due to the simultaneous (but unobserved) movement in a correlated input. Within the context of this thesis, correlations between input parameters, such as technology or resource costs in an [ESOM](#) are often used as a substitute for the description of the underlying mechanics. For example, the joint distribution over these input parameters could be better served by modelling the underlying phenomena, such as the process of learning.

One challenge in running a sensitivity analysis is that many model runs are required to cover as much of the input space as possible. With a computationally demanding model, there is a trade off between the time available, and the number of samples that can be run, and therefore the detail and resolution of the sensitivity analysis. Energy system models are relatively computationally demanding, with one run taking between 10 seconds and 10 minutes, depending on the size of the model. The ESME model, when optimising a multi-year solution which looks at the transition towards a low-carbon energy system from the present day, takes around 15 minutes to solve. There is also potential for the model to be infeasible (there is no optimal solution) at particular points in the parameter space. It is often not possible to determine this beforehand, so a sensitivity analysis technique which facilitates graceful handling of infeasible solutions is of benefit.

As energy system models are based upon linear programming, it is easy to obtain derivatives from the model solution which give a local and static indication of the sensitivity of an input parameter conditional upon the value of all the other inputs for that particular scenario. For example, the CPLEX solver includes a built-in routine which prints such information. As outlined in the previous section, derivatives can give misleading results if the model is of unknown linearity or is non-additive. In other words, this local sensitivity analysis approach will give only a shallow picture of the real sensitivity of the model parameters. However, there is scope for this ‘free’ sensitivity information to be used in conjunction with other techniques. This is left for future work.

One cannot make the assumption that the model is additive and linear in the input parameters, because there are a number of constraints, such as maximum build rates, that interact to cause non-linear behaviour. The portfolio of all technologies in the ETI-ESME model when taken together form a piecewise linear supply cost-curve. This curve represents the available supply in a given period of technologies at different price points. The width of the step in the cost-curve is parameterised by the maximum build rate for that period. However, while the portfolio of technologies that appear in the results of any particular model run *are* largely determined based on cost, other constraints also influence the combination of technologies chosen. For example, the effect of altering the cost of the technologies is to reshuffle the order of preference for the technologies. And due to the uneven piecewise steps, the resulting new installations will be affected in a non-linear manner. The direct effect on the total-system cost of displacing one technology for another may be marginal. However, the systemic effect of reshuffling the supply cost-curve is not immediately predictable and may have much larger consequences on the cost of the whole system³. Another example, is that the cost of mitigating CO₂ emissions tend to increase exponentially as the emission constraint increases (Usher et al., 2010b).

In summary, the characteristics of the ESME model suggest that the Method of Morris, discussed in the next section (3.3), is the most appropriate sensitivity analysis technique. The Method of Morris and extension outlined in Campolongo et al., (2007) indicate that the approach can feasibly handle the 200 inputs, of the ESME model, and by grouping similar inputs together, this can be extended to thousands of inputs. The minimal data requirements - ranges with a uniform distribution - are also a good match to that available for the ESME model. The ESME model in simulation mode uses a Monte Carlo sample generated with triangular distributions to represent the most-likely value, together with plausible upper and lower bounds. Many uncertain inputs are not included in the database of triangular distributions, but plausible ranges are relatively easy to find through a literature search. With grouped inputs, the Method of Morris scales according to $N(k + 1)$ where k is the number of groups, and N is the number of trajectories and is normally 10. Thus 100 input groups and $N = 10$ would result in 1100 runs, a feasible computation time for an ESOM. Branger et al., (2015) describe how the Method of Morris is resistant to model crashes (infeasible solutions) as it is possible to remove from the analysis those runs where the crashes occurred without affected the rest of the analysis. Finally, the Method of Mor-

³ The papers of Kesicki, (2012) and Kesicki et al., (2011a,b, 2012) illustrate some of the issues of related marginal abatement cost (MAC) curves within the context of a whole-system modelling

ris is a [global sensitivity analysis](#) approach and deals well with non-linear and non-additive models.

3.3 EXTENDING THE METHOD OF MORRIS TO COMPUTE OPTIMAL INPUT SAMPLES

In this section, I introduce an extension to the Method of Morris building upon work started in Campolongo et al., (2007). With computationally intensive models, such as [ESME](#), the number of simulations which can be feasibly performed is a limiting factor. With a small number of simulations, there is a risk that the pseudo-random sampling procedure used in the Method of Morris creates a biased sample which in turn biases the results. In this section I show how a combinatorial optimisation procedure can be used to pick the most-different [trajectories](#) from a pool of [trajectories](#) to create an optimal sample. After this introduction, the technical details of this optimisation procedure are introduced in Section 3.3.1. I then discuss the effect of the selection of probability distributions for the inputs, grouping input parameters and the open-source software package in which these procedures are implemented.

The Method of Morris (Morris, 1991) is a simple average of derivatives over the space of input parameters. The advantage of the approach is that it gives a reliable sensitivity measure for relatively large numbers of input parameters (10s to 100s) with a minimum of model runs. Thus the model is explored over a wider combination of input values than gradient-based local approaches. An outline of the sampling procedure and method for computing the elementary effects and subsequent sensitivity measures are described in Appendix A. One advantage over alternative techniques in which input parameters are grouped, such as fractional factorial approaches (see Saltelli et al., 2008b), is that each input parameter is permuted and examined individually, so cases in which two influential parameters cancel one another out are revealed.

However, the effectiveness of the Morris technique for estimating elementary effects relies upon a sampling method that gives a statistically equal weight across the possible values of each individual input parameter. The input space for each parameter is divided into a discrete (and even) number of [levels](#) of equal size. Ideally, histogram plots of the resulting simulations should show an uniformly distributed sample for each variable. If the sample is biased, then the elementary effects for parameters which are non-linear or interact with other parameters would also be biased, giving misleading results.

There are two main input parameters to set up the sample. These include the number of [levels](#) p into which each parameter input range is divided, and the number of [trajectories](#) N . These parameters are

related. As shown in Figure 2, the distribution of the input sample improves as the number of trajectories N increases.

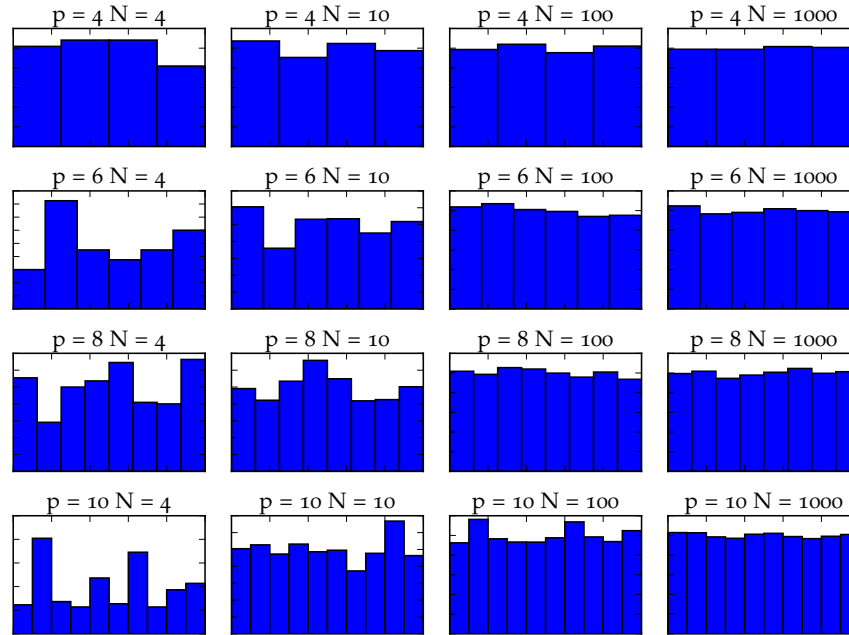


Figure 2: This figure shows the effect upon the randomly generated input sample to a model of the number of levels p as the number of trajectories N increases.

The number of samples N needs to be as high as possible, but it is also important that the value of N isn't too low compared to the number of levels p that are chosen. For instance, Saltelli et al., (2008a) suggest that $N \geq 10$ when $p = 4$. As such, little is gained from defining a high resolution analysis with a big value for p , as a large number of trajectories N will be required to give an even sample distribution across all parameters and parameter values. And Wainwright et al., (2014) found that using a small value for p reduces the likelihood of under-estimating μ - the mean elementary effect of an input, lending extra weight to argument that p should be kept as small as possible. This has the advantage of reducing the number of trajectories N that are required.

For computationally demanding models such as ESME, N has to be in the lower 10s or 100s for the approach to be feasible. However, given the quasi-random approach taken to generate the sample trajectories, there is a small chance that a very poorly distributed sample is generated, particularly if the number of trajectories N is very small. This would give a correspondingly poor estimate of the elementary effects and misleading results.

To avoid this, Campolongo et al., (2007) proposed a method that obtains an optimal combination of trajectories from a large pool of quasi-randomly generated trajectories. Ruano et al., (2012) note that

this brute-force method requires a huge computational effort, and suggest a heuristic-based technique to compute a local optimum for analysis of models with many input parameters. They claim that the approach dramatically reduces the time needed to obtain an input sample which is ‘good enough’. However, they are only able to compare the effectiveness of their approach with the brute-force method with a small N . There could be an advantage to using a combinatorial optimisation approach to determine the globally optimal solution for medium values of ($10 \geq N \geq 100$) which is quicker than the brute force method. This latter approach is described and formulated in the next section.

3.3.1 *Optimal Trajectories*

Given the computational demands of the ETI-ESME model, and the large number of inputs, it is imperative that the sample size N is kept as small as possible, while ensuring that the sample is fair and unbiased. While the sample shown in Figure 2 for $N = 1000$ is near perfectly uniform for each of the variables⁴, the number of simulations required to achieve this is effectively computationally impossible. For groups of variables numbering up to 100, a value of N around 10 to 20 is feasible. The upper four rows of histograms shown in Figure 3 for $N = 10$ are indicative of the magnitude of bias inherent in this approach (it is far from uniform for each variable). This section describes our approach for obtaining a better sample for low values of N .

Campolongo et al., (2007) extended the work of Morris, (1991), by computing optimal combinations of *trajectories* to avoid situations in which the randomly generated *trajectories* fail to evenly cover the input space and to improve the quality of the estimate of the total sensitivity index provided by the elementary effects method. Below we introduce a further extension, which allows the computation of optimal *trajectories* using a binary integer programme instead of the brute force methods used by Campolongo et al., (2007).

The rationale for computing optimal *trajectories* as an optimisation problem is that the size of the brute force problem increases exponentially mainly as a function of a N , the number of initial *trajectories*, and as o the number of optimal *trajectories* required is typically much smaller than N . This is because after computing a simple Euclidean distance between each pair of *trajectories*, the sum of each possible combination of *trajectories* is computed. The combination with the highest score is that with the most different *trajectories*.

⁴ Note that this plot shows all of the variables in each histogram, so that it is not possible to directly compare with Figure 3

First a distance matrix d_{ml} is computed using a Euclidean distance:

$$d_{ml} = \begin{cases} \sum_{i=1}^{k+1} \sum_{j=1}^{k+1} \sqrt{\sum_{z=1}^k [X_z^{(i)}(m) - X_z^{(j)}(l)]^2} & m < l \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where k is the number of input parameters and $X_z^{(i)}(m)$ indicates the z th coordinate of the i th point of the m th trajectory as described in Saltelli et al., (2008a). Thus, the distance matrix d_{ml} is the sum of the distances between each pair of points in the trajectories under examination. This distance matrix is then used to compute a distance measure D^2 which is the square root of the sum of each squared 2-subset for each possible combination of trajectories:

$$D_c^2 = \sqrt{\sum_i^I \sum_{j=i+1}^I d_{ij}^2} \quad \forall (i, j) \in H_c \quad (7)$$

where $c = 1, \dots, \binom{M}{k}$, H_c is a unique vector of combinations of trajectory indices for each combination c . For example, to determine the highest scoring combination of 4 trajectories from a pool of 10 possible trajectories, $\binom{10}{4} = 210$ separate computations would need to be performed, each of which requires the summation of $\binom{4}{2}$ distance measures. At the magnitudes suggested by Campolongo et al., (2007) ($N \sim 500 \rightarrow 1000, r = 10, 20$), 10^{23} to 10^{41} computations would need to be performed to find the maximally scoring combination of trajectories. Clearly this is unrealistic.

When formulated as an optimisation problem, the problem size increases proportional to the square of N , while o does not affect solution time. The decision variables x_{ml} increase with the size of the distance matrix, while the constraints are also generated for each trajectory. The distance matrix takes the form of a lower-triangular matrix, and has several attributes for formulating the optimisation problem which are discussed below.

The optimisation problem is to maximise the score of a combination of o trajectories from the pool of available (randomly generated) trajectories N .

$$\max \sqrt{\sum_{i=1}^I \sum_{j=1}^J (d_{ij})^2 x_{ij}} \tag{8}$$

$$\text{s.t.} \quad \sum_{j=1}^J x_j \leq k \tag{9}$$

$$d_{ij} \leq x_i \quad \forall i, j \tag{10}$$

$$d_{ij} \leq x_j \quad \forall i, j \tag{11}$$

$$x_i + x_j \leq 1 + d_{ij} \quad \forall i, j \tag{12}$$

$$\sum_{i=1}^I \sum_{j=1}^J d_{ij} \leq k(k-1)/2 \tag{13}$$

$$x_j, d_{ij} \in [0, 1] \quad \forall i, j \tag{14}$$

$$\tag{15}$$

The intuition underlying the formulation of the constraints is as follows. The distance matrix is triangular, with the axes representing the index of the trajectory. If two scores are picked e.g. $x_{1,2}$ and $x_{2,5}$, then there are a finite number of trajectories that are legal combinations with these two trajectories within the confines of the distance matrix. The third score must share at least one of each of the other score's trajectories. So the only possible choice to match the example scores picked would be $x_{1,5}$.

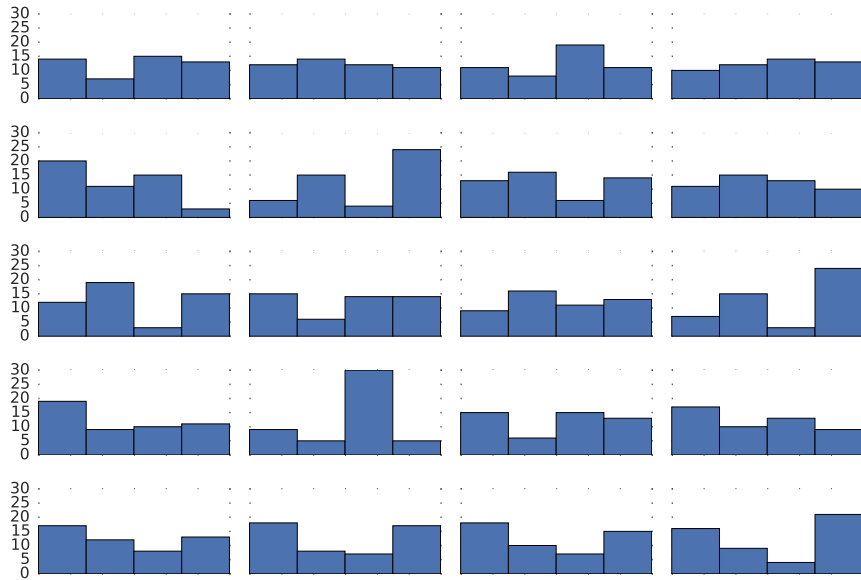


Figure 3: This figure shows the effect upon the input sample to a model, without (top four rows) and with (bottom) the optimisation procedure. The optimisation avoids the worst incidences of sampling bias, such as for the second variable in the fourth row.

Figure 3 shows the difference between optimised and quasi-random generation of the samples. Each of the first four rows is an equal segment of the sample from which the optimised *trajectories* in the fifth row were produced. Each column represents one of four variables. Each histogram shows the frequency distribution of the *levels* of the variable over ten *trajectories*. Following the second column from the first to fourth rows, we can observe that within the quasi-randomly generated *trajectories* there is a large capacity for change in the frequency with which individual parameter *levels* are sampled. In contrast, the first row, representing the frequency distribution of the *levels* for the four variables in the first ten *trajectories* in the sample bundle, seems to cover the variables in a uniform fashion. Figures 48 and 49 showcase the two extremes. Figure 48 demonstrates that the best combination of the same *trajectories* are a qualitative improvement, and also that there is an unavoidable element of variation between the optimal combinations due to the variation in the pool of *trajectories*. In the latter case, the figure shows what would happen if the worst combination of *trajectories* were combined into a sample. Generally, the frequency distributions for each variable in the best case (Figure 48) are closer to uniform than the worst case samples.

3.3.2 *SALib: A Python library for conducting sensitivity analysis*

The extension to the Method of Morris was implemented within the existing open-source Python library called SALib (Herman et al., 2015). The codebase of the library is hosted on <http://www.github.com/SALib/SALib> and the version control software allows the contribution by the author (@willu47) to be viewed in context. The optimisation code was written in gurobipy, and the optimisation problems solved using the commercial Gurobi solver. The code is completely open-source and available for download and use under an MIT license.

3.3.3 *Rationale for selecting distributions*

The *global sensitivity analysis* uses uniform distributions for all inputs. The rationale for using uniform distributions over any other is described below.

When conducting a sensitivity analysis of a model for which the range of values are well known (i.e. in the quadrant of *Knighian Risk* in Appendix C), then the choice of distribution is relatively straightforward. The same choice of distributions when dealing with a model that crosses several of the quadrants of uncertainty, such as an energy system model, is not as straightforward.

A uniform distribution presents hard upper and lower bounds to the range of possible values that a parameter can take, and implies indifference as to the likelihood of the values between these bounds.

While the existence of hard upper and lower bounds is hard to justify, for a screening analysis, these bounds are merely taken as the highest and lowest values within which the parameter is perturbed. A reasonably liberal approach can therefore be taken, right to the bounds of plausibility as the intention is to test the model at extremes as well as within more normal operating ranges. The Method of Morris subsequently picks a discrete number of ‘levels’ (usually four) between the upper and lower bounds, and with a uniform distribution, these will be of equal distance from one another and the bounds. Given the four levels over which input values are sampled, the exact distribution is less important than the absolute range over which the parameters are varied.

The effect upon the sampling procedure of a normal over a uniform distribution is to squeeze the central two values according to the shape of the normal distribution. The resulting mean elementary effect would be mildly reduced (depending upon the exact role of the parameter in the model), with the uniform distribution giving in most cases the worst case value. This effect is not pronounced when sampling only uses four levels, but would become more significant as the number of levels is increased. This is because central values would be given more weight in the averaging of the elementary effects. Because of this, increasing levels above four is more likely to underestimate the sensitivity of a variable Wainwright et al., (2014). In some cases, the use of uniform distributions could result in the ranking of the elementary effects to be less accurate than if the true distribution were known, but even then only if the distribution were radically skewed, or extremely narrow. However, I maintain that this inaccuracy is likely to be far less than that introduced through assigning distributions where they are not warranted.

Finally, the assumption of uniform inputs is conservative. In the worst case, if the assumption is dramatically wrong and the true distribution is strongly skewed, a Type I error will occur; wrongly identifying an unimportant input as important. This is preferable to a Type II error where an important input is misdiagnosed as unimportant. However, there is a risk that if the upper and lower bounds of the uniform distribution are substantially narrower than those of the true distribution, than a Type II error could occur. It is strongly suggested that the bounds are estimates as wide as is deemed reasonable.

The upshot is that the information on elementary effects obtained from the Morris analysis is affected by the number of levels chosen, and that there are interactions between the number of levels, the importance of selecting the ‘correct’ distribution, and the accuracy of the results. There is a small chance that quality of the sensitivity measures suffer in situations where the true distribution of a parameter is strongly skewed or exhibits substantial kurtosis. However, for many of the parameters the pertinent information from the Morris screen-

ing analysis is crude enough that many of these issues will wash out, while still giving useful insights into the behaviour of the model.

3.3.4 Grouping input parameters

Parameter groups were chosen according to the similarity of the parameters to one another in the [ETI-ESME](#) model. The selection of groups is not critically important as the grouping of parameters does not unduly affect the results of the Morris screening analysis for individual input parameters, other than to aggregate them together.

In many cases, there are parameters that represent logical groups due to their similarity in function across one demand sector of a technology type. For example, in the [ESME](#) model analysis presented in [Section 4](#), all district heating network technology cost parameters were placed within a group. These parameters are likely to move together (be highly correlated) in reality, and play a small role in determining the cost of the energy system.

An early experiment in which all liquid fuels (such as gasoline, diesel and aviation fuel) were grouped, contained too much of the variance of the model output to be a sensible grouping of parameters. These important parameters were subsequently treated independently from one another, despite their likely correlation in a real-world situation.

To reiterate, grouping parameters because of their likely correlation is not suggested or necessary. When grouped, the amended Morris sampling method moves each of the parameters within the same group at the same time in each line of a [trajectory](#). However, this movement can occur from different starting points, and in different directions. Elementary effects are then computed for each parameter individually, μ^* computed and finally averaged over the group to give a measure for the whole group. The grouped measure of influence is therefore a sum of the influence of each parameter individually.

This technique for the grouping of input parameters thus avoids issues whereby a un-influential parameters are incorrectly identified as influential, because they are in a group with an influential parameter, and *vice versa*. The [Appendix A](#) shows that while μ^* (mean of the absolute elementary effect) is a valid metric when using groups of parameters, the measure of standard deviation σ is not usable, as this relies upon the metric μ which it is not possible to compute when using groups. The error bars on the plots in the results in [Chapter 4](#) show the 95% confidence interval instead.

The limitation of grouping input parameters is that the measure of non-linearity/interaction σ is unavailable, while the benefit is that far fewer model runs are required. Ideally, unimportant parameters may be grouped together, to allow σ to be computed for the more important parameters.

3.4 INTEGRATING THE SENSITIVITY ANALYSIS, LEARNING MODEL AND ETI-ESME

The methodology and results of the sensitivity analysis approach outlined in this and the next chapter are interesting in isolation, and have a number of methodological insights for developers and users of such models. However, the fundamental aim of the sensitivity analysis is to inform the selection of *influential* uncertainties amenable for analysis in the learning model described in Chapter 5. In this setting, there are a number of challenges to overcome. These include how uncertain inputs are translated into inter-temporal trends, the link between uncertain inputs and the interpretation of the global sensitivity analysis results in the context of the learning model. These issues are outlined below.

3.4.1 Mechanism for treating the variation of inputs over time

Uncertain inputs in the ETI-ESME model are modified using a system of indices. The values for the parameters in years other than the start year (2010) and end year (2050) are computed from a linear interpolation of the 2010 and 2050 values, multiplied by a scaling factor (see Equation 16). This enables the *trajectory* of a parameter over the period 2010 to 2050 to take a predefined shape, and this shape is then permuted up or down depending upon the value of the 2050 value.

$$C_{t,k}^{Cap} = C_{2010,k}^{Cap} \cdot \sum_{ix,k} w_{ix,t,k} \cdot s_{ix,t,k} \quad (16)$$

such that

$$\sum_{ix} w_{ix,k} = 1 \quad \forall \quad k \in K \quad (17)$$

where

$w_{ix,k}$ is the weighting assigned to the indices enabling indices to be ‘blended’ for index ix and technology k

$s_{ix,t,k}$ is the index scalar dictating the *shape* of the cost *trajectory* for index ix , technology k in time period t

During selection of the random values in a Monte Carlo simulation or sensitivity analysis, the values for the year 2050 are drawn from a sample process (Equation 18). These are then used to rescale the values between 2010 and 2050.

$$C_{s,t,k}^{Cap} = C_{t,k}^{Cap} \times \omega_{s,k} \quad (18)$$

where

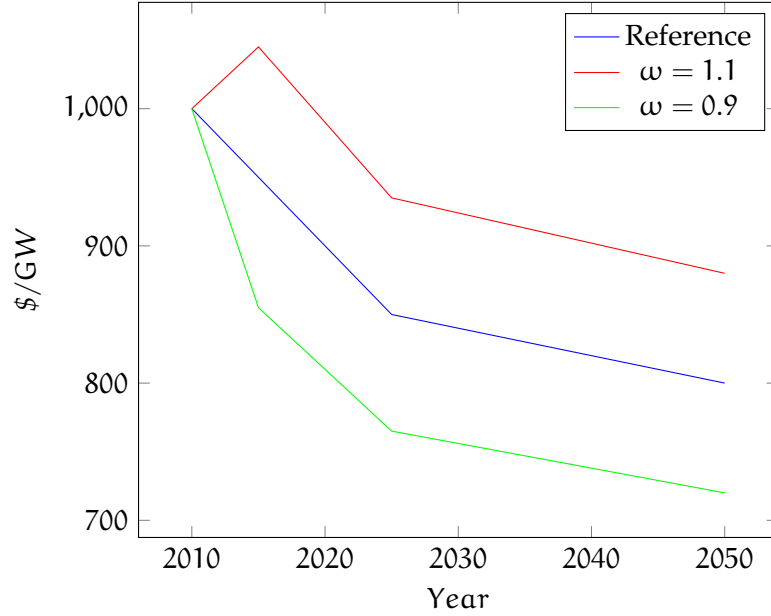


Figure 4: Example simulation capital cost trajectory for different samples of ω

$C_{s,t,k}^{\text{Cap}}$ is the capital cost for technology k in simulation s in time period t

$\omega_{s,k}$ is the random value in simulation s for technology k

Thus the values for the inputs prior to 2050 are a function of the product of the simulated 2050 value and the weights assigned to denote the trajectory shape. However, these linearly interpolated inputs may have an important role in defining the model output. In some cases, the 2050 value may have no direct effect upon the output, and rather exerts influence through its role in defining the value of the linearly interpolated input values. This assumption makes it more complicated to extrapolate the influence of any other input upon the output than those expressly investigated in this work. The obvious solution to this, would be to expressly include these indices in the sensitivity analysis. However, this greatly increases the size of the analysis. While the grouping of inputs could mitigate this increase, a second issue is that independently varying prices/costs/efficiencies in each year could result in nonsensical inputs such volatile increases and decreases in a technology cost, where a long-term decrease may be more likely.

In summary, the assumption made in the model formulation, that inputs follow a linearly interpolated increase or decrease from the value in the model start year is a useful one in terms of reducing the degrees of freedom of the model. However unrealistic this assumption is, the alternatives could also result in unrealistic parameter trajectories. What is most critical is whether these assumptions are acceptable for the purposes of identifying the most influential input

parameters. The answer is that the temporal influence of individual inputs is not particularly well explored, but that the simple increase/decrease in 2050, also leads to an increase/decrease in 2030 (but of a smaller magnitude).

An alternative (and more sophisticated) approach is taken in Anadon et al., (2011, §2.5.3.3) in which longitudinal dependencies are elicited to infer the degree to which 2010 and 2030 technology costs inform 2050 technology costs. In contrast, the ESME model does not use longitudinal dependencies, assuming a fixed correlation over time regardless of the simulated value.

3.4.2 Mapping input uncertainties to agency

A key link in the chain between the input parameters to the ESME model, and the learning model described in Chapter 5 is the nature of the uncertainty surrounding the ESME inputs. The discussion in Appendix C distinguishes between *aleatory uncertainty* and *epistemic uncertainty*. In short, the former is irreducible and can be managed using risk-based approaches, while the latter is amenable to learning, waiting or research.

However, the ability of a decision maker to take action, (their *agency*) dictates the portfolio of approaches that are possible. If action is possible, then the decision maker's ability to control or influence the resolution of uncertainty is restricted by the nature of uncertainty. The actor based approach described in Hughes et al., (2012) reiterates how the decision to take action can itself increase or reduce uncertainty.

Table 6 gives a range of examples of input parameters to the ESME model, the corresponding nature of the uncertainty, and the appropriate methodology given this nature. *Hedging* refers to a risk-based approach, where probability weighted outcomes are used to decide upon an optimal near-term strategy. *Learning* refers to the decision to wait for more information, or conduct an examination of the uncertainty. *R&D*, or research and development is a special case of learning, where a financial investment is made in the interest of resolving a specific uncertainty.

Aleatory uncertainties cannot modelled using a decision dependent approach. This is because an aleatory uncertainty cannot be further reduced, and so to invest in a learning process that will not generate new information is self-defeating. Epistemic uncertainty, by definition, may be reduced.

Note that mature technologies and resources, for example a coal-fired power station, and crude oil respectively, are unlikely to be amenable to an R&D approach. The uncertainty surrounding crude oil prices is largely aleatory, and no amount of expenditure on R&D is likely to reduce the volatility of oil prices, or conversely, increase

our knowledge about how the oil price will evolve⁵. On the other hand, oil extraction costs are less subject to aleatory uncertainty and perhaps more exposed to epistemic uncertainty. While in [ESOMs](#) oil prices are generally more relevant than extraction costs, energy system models have a mix of both costs and prices ⁶.

Table 6: Mapping uncertain inputs to agency. DD indicates whether the uncertainty can be modelled as decision dependent - using the learning model in Chapter 5.

Parameter	Uncertainty	Methodology	DD
Fossil Resource Prices	Aleatory	Hedging	No
Novel Resource Price	Aleatory/Epistemic	Learning/R&D	Yes
Max Resource Quantities	Epistemic	Learning	Yes
Mature Tech. Capital Cost	Epistemic	R&D	Yes
Novel Tech. Capital Cost	Epistemic	R&D	Yes
Novel Tech. Avail.	Epistemic	R&D	Yes
Novel Tech. Max. Build Rate	Epistemic	Learning	Yes
Mature Tech. Avail.	Aleatory	Hedging	No

On the other hand, a novel resource, such as domestic biomass, is subject to a mixture of both aleatory and epistemic uncertainties. For example, the production costs and available quantities of the UK's indigenous biomass resource are not well understood, and there is scope for conducting research projects to better explore the interactions between biomass crops and the food and water systems in the UK ([Bajželj et al., 2014](#); [Bazilian et al., 2011](#)). Even long-run production costs for biomass are subject to a mixture of aleatory and epistemic uncertainties, especially as climate impacts upon the biosphere are still relatively unknown, and natural variability from decade to decade can influence the supply (and thus the market price) of crops⁷.

The maximum rate at which technologies may be rolled out is of potentially systemic importance. Constraints on the rate of growth or a maximum capacity can prevent the construction of the economically optimal quantity of a technology, forcing investment in more expensive technologies. However, such constraints may not be apparent until a particular technology commences construction, or may be

⁵ Note that from a Knightian Risk perspective, these two examples are identical

⁶ For example, in the [ESME](#) model, costs and prices are confused, with Biomass Resource Cost of Biomass Imports actually representing the market price of biomass available to the model, while domestic biomass resource costs effectively represent the cost of domestic production.

⁷ ESME includes constraints on availability, rather than using supply-cost curves to represent increasing marginal prices of resources as in other energy system models such as [TIMES](#).

dependent upon an otherwise unrelated infrastructure outside the scope of the model.

Clearly, there are limits to our ability to reduce the uncertainty associated with various inputs to the [ESME](#) model. These limitations derive from the nature of the uncertainty, and also from the constraints that apply to deciding to resolve an uncertainty. Depending upon the results from the sensitivity analysis, it may be possible to ignore the less influential input parameters, focusing on just those uncertainties that determine the bulk of the change to the energy system. Still, there is a complex relationship between those uncertainties which must be managed, those that are amenable to learning, and those that can safely be ignored, and the sensitivity analysis can enable such relationships to be identified through the indication of interaction effects.

3.4.3 *Identifying meaningful output metrics*

As an optimisation model, the [ESME](#) model minimises total energy system cost. This is the discounted sum of total capital, fixed and variable costs associated with investment and operation of technologies across the whole energy system (Equations 1-2). This includes costs for new buildings (domestic and public) and the total investments for transport technologies (including private passenger vehicles such as cars). Total energy system cost is therefore the most obvious output to use in the sensitivity analysis, particularly as this is used to derive a revenue function for the learning model described in Chapter 6. This would indicate which inputs most influence the objective function, and are thus likely to result in changes to the model solution. However, the insights that modellers obtain from energy system models are often quite different to a narrow focus on the cost of the energy system. While, the objective function is the mechanism by which the energy system model determines which package of decisions is optimal, two otherwise radically different energy systems can be very similar in cost.

The nature of the response of an output metric to the movement of an input variable could offer useful information for modellers. For example, a large linear response to an input variable is likely to indicate an over-constrained model, or lack of options for technology or resource substitution. Non-linear responses and interaction effects are equally important to identify, because they could help identify clusters of variables that together drive the energy system in a (un)desirable direction.

Kriegler et al., (2015) use high level indicators obtained from combinations of outputs from integrated assessment models which allow comparison between the models. The indicators mirror the main insights that are derived from the model responses to carbon price signals. The indicators include carbon intensity, energy intensity, and

the extent of the change in the energy system. While the [ETI-ESME](#) model is smaller in scale, and narrower in scope than the integrated assessment models in this study, the basic idea of using a quantitative indicator to match what a model user may wish to obtain from the model is a useful one. In this vein, [Table 7](#) lists the model indicators which appear in the results chapter ([Chapter 4](#)).

Table 7: Model output indicators

Indicator	Unit
Discounted total energy system cost	£bn
Carbon price	£/tCO ₂ e
Renewable % of electricity capacity	%
Electricity Production	TWh

The objective function of the [ESME](#) model minimizes the sum of individual representative years and does not provide an objective function which is interpolated between modelled time periods. As such, the objective function must be post-processed — scaled by a factor to represent the discounted sum of the intervening years between time periods. Due to computational constraints and issues with the stability of the model using five-year time-periods, the sensitivity analysis was performed using 10-year time-periods. The objective function was subsequently scaled by a factor calculated as follows $\sum_{i=0}^{i=9} (1 - 0.035)^i = 8.56$ (using a discount rate of 3.5%).

3.4.4 *Running a Morris analysis on an optimisation model*

I conclude with a discussion of some of the potential issues that may be faced when using the Method of Morris with optimisation models.

While the Method of Morris provides a robust sensitivity measure, when applied to optimisation models there are a number of potential issues that can occur. These could potentially bias the results. However, sometimes even the failure of the model gives new and useful information. Some common issues are discussed below.

3.4.4.1 *Infeasible solutions*

It is quite possible for a particular combination of inputs to result in an infeasible solution to an optimisation model ⁸. This results in missing, or worse, spurious data populating the results from which the elementary effects are computed. Generally, these can only be

⁸ Note that infeasibility is most often caused by the imposition of conflicting constraints, or the triggering of a constraint by an extreme coefficient value, whereas extreme values of coefficients alone will give just extreme results, rather than infeasible solutions

identified manually. In the case of the ETI-ESME model, it was easy to identify missing data, as the database contained either a zero, or a very high value in the results. When plotted unmodified, this revealed an extremely large confidence interval across multiple input parameters. Branger et al., (2015) suggest removing just those elementary effect calculations corresponding to the missing data. Each infeasible simulation corresponds to two elementary effect calculations, fewer $(n + 2)$ if n simulations are consecutive. Alternatively, the data can be replaced with the mean value of the sample for that parameter, which would effectively skip the computation of the elementary effect.

From a modelling perspective, the investigation of infeasible solutions is a useful stress-testing of the model formulation. An infeasible solution indicates that the combinations of inputs for which the model is ill posed, and can indicate a structural weakness or bug in the model formulation. Thus a useful side-effect of the Method of Morris is that the model is stretched outside of the usual operating boundaries. The input sample generating using the method in Appendix A ensures that the value of just one parameter changes from one simulation to the next (within each trajectory). This means that a simple factor mapping can be conducted *ex post* where ranges within the input space can be linked to infeasible solutions.

3.4.4.2 Lack of results data when results are too disaggregated

In circumstances when an output metric is selected which does not show enough variation given the size of the simulation run, the sensitivity measures will not be reliable. This is indicated in the results through large confidence intervals, relative to the value of μ^* . The outputs from deterministic optimisation models can be particularly prone to this behaviour when used with the Method of Morris, because an output may vary only between four levels or fewer levels as a function of the subset of the input parameters which are also only varied over four discrete levels. As a consequence, the approach works best when using more aggregated continuous output metrics, such as measures of cost, prices and capacity.

An example of this behaviour can be seen in Figure 9 on page 74. In this case, the LGV Capex and Demand Capex parameters vary over four levels while the computed output metric (percentage of renewable electricity capacity) varies over fewer levels, or stays largely stationary with a few large divergences. Solutions for this instance could include running more simulations, using more than four levels in the input sample, including more inputs in the global sensitivity analysis or choosing a different output metric.

3.5 SUMMARY

In this chapter, we've explored the [ESME](#) model, an [ESOM](#) of the United Kingdom. Given the fundamental characteristics of the [ESOM](#) - a large number of inputs, data limitations, computational demands and non-linear outputs - the Method of Morris was selected as the most appropriate sensitivity analysis approach. An extension to the Method of Morris was proposed, which computes the globally optimal combination of randomly generated [trajectories](#). This ensures that the sensitivity analysis input sample to the model is fair and unbiased, giving good coverage across the input space. This is particularly important given the small number of simulations (and therefore [trajectories](#)) that it is computationally feasible to perform on the ESME model. Finally, a set of challenges were highlighted, including the mechanism for treating dynamic uncertainty, building a framework to marry agency, influence and uncertainty, identifying meaningful output metrics, and the soft-linking of a perfect foresight model with a learning model.

This chapter is divided into four sections, describing the results of the global sensitivity analysis of the [ESME](#) model. After an overview of the results, the 119 inputs of the model (or 200 total uncertain parameters) that were included in the analysis are investigated in separate categories, and analysed across the output parameters outlined in [Section 3.4.3](#). The output metrics were chosen to represent the different insights of the energy system model that were likely to be used by modellers. However, the main output parameter of interest for this thesis is the revenue function of the learning model (a description of which follows in [Chapter 5](#)). The revenue function is derived from the total cost, which is the discounted sum of capital and operation costs across the energy system (shown in [Equations 1 to 5](#)).

In [Section 4.1](#), I give an overview of the sensitivity analysis results. In the following [Section 4.2](#), I investigate the input parameters according to their categories, such as resource constraints and technology costs. In [Section 4.3](#), the results from a side experiment, in which the sensitivity of the elastic demand parameters were investigated, to compare with the findings from [Pye et al., \(2014\)](#). Finally, I discuss the sensitivity analysis results and compare the efficacy of the approach of managing uncertainty to that of scenario analysis, Monte Carlo sampling and exploratory modelling and then identify some policy implications of this work.

4.1 RESULTS FROM THE MORRIS SCREENING: OVERVIEW

In this section, I first reiterate the sensitivity analysis metrics used over the following pages. I then present the results of the sensitivity analysis.

4.1.1 *Interpreting the results*

The sensitivity analyses were conducted using the Method of Morris, grouping input parameters, computing the elementary effects, and finally calculating the metric μ^* as described in [Appendix A](#). The grouping of the input parameters is detailed in [Appendix D](#). The metric μ^* corresponds to the group average of the mean of the absolute elementary effect caused by the movement of the group of input parameters over their range. μ^* is therefore a proxy for the mean of the distribution of elementary effects, while the 95% confidence interval obtained through bootstrapping gives an impression of the deviation

around the mean. The units of μ^* are the same as for the output upon which μ^* is based.

Within a group, each of the inputs moves independently, and so the grouped value of μ^* gives the sum of total influence caused by the group of input parameters within that group, given their independent behaviour within their ranges. The computation of the grouped μ^* is done in such a way that interaction effects are explicitly managed. For example a dominant input parameter within a group will not cancel out interactions in the opposite direction from the remainder in the group. The grouping of inputs for the purpose of the sensitivity analysis still allows the inputs to move individually and so is distinct from the aggregation of model inputs under one *dummy input*.

While the advantage of grouping the inputs means that many more inputs may be incorporated into the analysis with comparatively little computational penalty, the grouping does mean that the metric σ may not be used to quantify the standard deviation of the elementary effects. In the non-grouped version of the Method of Morris, σ gives a measure of the interaction effects or non-linearity of the input upon the output. With the grouped method, only the bootstrapped confidence interval may be used which has a different interpretation to that of σ . The confidence interval (CI) signifies the interval in which in 95% of the times the value for μ^* will lie if the sensitivity analysis is repeated. Unfortunately, no information regarding the non-linearity and interaction effects may be obtained from the use of the confidence interval. The confidence interval does indicate when there are too few runs to estimate the sensitivity of a parameter with the desired precision. As the number of runs increases, the expectation is that the range covered by the confidence interval would decrease.

To test the effect of grouping the input parameters, I ran a more detailed screening analysis, which fixed the least sensitive parameters from the initial screening to their mean values, and varied the remaining ungrouped parameters over 8 [levels](#) using 20 [trajectories](#) of 30 model runs giving 600 simulations in total. In particular, the group Car ICE was disaggregated into its component parameters Car ICE A/B Capex, Car ICE C/D Capex etc. The ranking of the most sensitive parameters remained among the most influential parameters of those in the more detailed analysis. This confirmed that the fixing of the least influential parameters does not materially affect the variation in the total energy system cost as would be expected.

The plots that follow do present a value for σ whenever an ungrouped input parameter is presented alongside a group of inputs. For example, the parameter Biomass Max Resource Qty only contains the parameter Max Resource Quantity for the Biomass product. In this case, values for μ can also be used to obtain extra information such as the direction of the effect on the output. The diagonal lines emanating from the origin on the plots represent different gradients

of the coefficient of variation (CV) which shows the variability in relation to the mean. CV is calculated from σ/μ^* . Broadly speaking, $CV > 1$ indicates a highly interactive or non-linear parameter, $0.5 > CV > 1$ indicates a moderately interacting or non-linear parameter, with the remaining parameters near-linear in effect or non-interacting. The lines divide the plot into areas to allow easy identification of the parameter behaviour.

4.1.2 Overview of the results

Unless specified, the following results were run under an 80% reduction in CO₂ emissions below 1990 levels.

In Figure 5 the values of μ^* are shown for the most influential groups of inputs for the total cost of the energy system (the objective function of the ESME model). The Figure shows 20 of the 31 tested groups of the 119 input parameters (of the 200 total available uncertain parameters) to which the ETI-ESME model is most sensitive ranked in order of influence, the remainder having little or no influence upon the output of the model. A list of all the groups is contained in Appendix D.

The influential parameters are clearly identifiable, with a very few parameters responsible for the majority of the variation in total energy system cost. The top five groups of inputs responsible for the majority of the variation in total system cost, include two of the three fossil fuel resource costs, a constraint on domestic biomass availability and aggregate build rate constraints for both CCS technologies and the electricity sector.

The value of σ available for the resource cost parameters show that they do not interact much with other parameters. Liquid fuel is almost all consumed by transport, and natural gas by the electricity and residential heating sectors. Both of these uses are relatively consistent over all levels of CO₂ mitigation targets or their effects are largely linear. The model is constrained so that liquid fuel and gas is used at any cost, and near-term behaviour consumption of liquid fuel and gas is locked into the system due to legacy investment decisions and the existing stock of technologies. Changes in the total energy system cost are almost a linear function of the cost of these resources. Changes in the costs of fossil fuels do not affect the model decisions. For example a lower fossil fuel cost does not result in a significant increase in the demand for the fuels, because of the constraint on emissions.

The values of μ^* give some indication of the magnitude of variation of the total cost of the energy system caused by the movement of the input (or groups of inputs) over its (their) defined range. The top four input parameters each cause a 3-4% variation in the total cost. These are not necessarily additive - there may be interaction effects, but it is

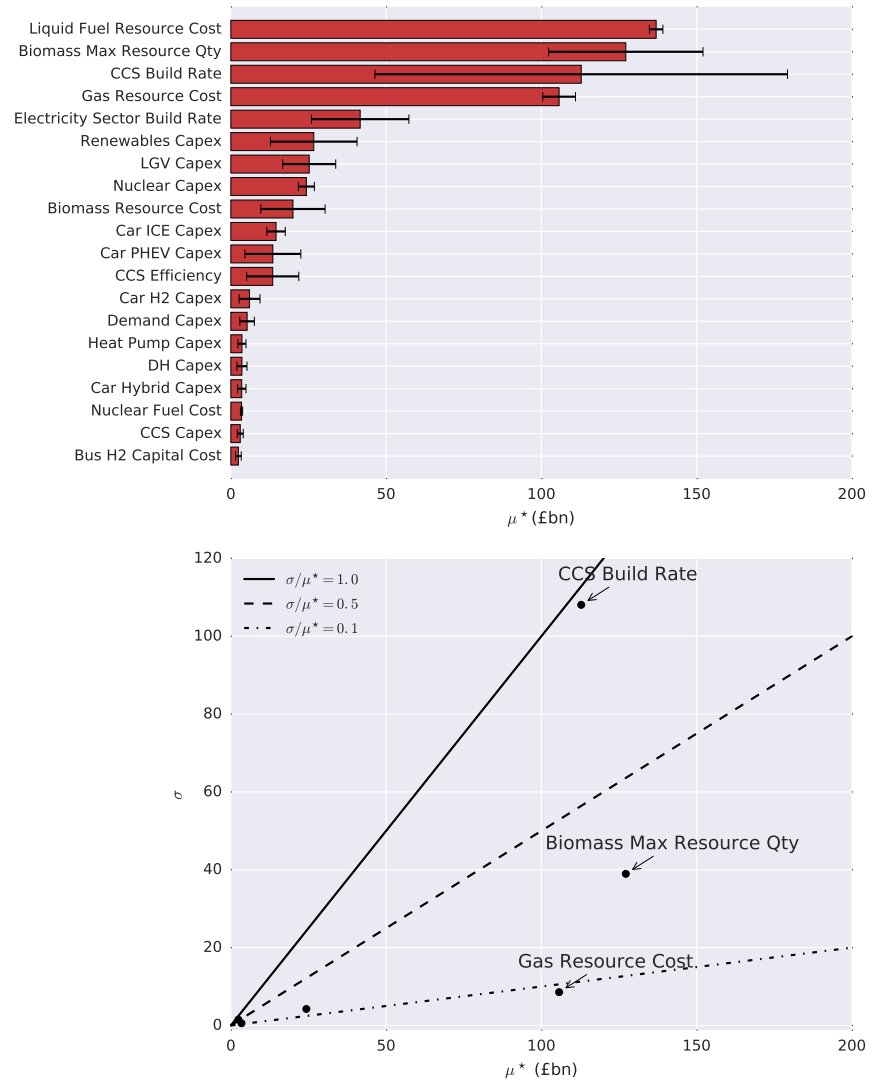


Figure 5: Influence on total cost of 31 inputs groups of parameters

not possible to understand these from the confidence intervals. Considering that the objective function comprises a considerable fraction of the UK economy, a potential 16% variance seems comparatively large.

Figure 6 shows the results for the marginal price of carbon dioxide (CO₂) emissions in 2050. Under an 80% emissions reduction target below 1990 levels, carbon emissions must decline to less than 118MtCO₂/year by 2050. The marginal price of CO₂ is a function of the multi-dimensional supply-cost curve, itself a function of the technologies and constraints defined within the model database, together with the operating and emission constraints. Policy makers interpret

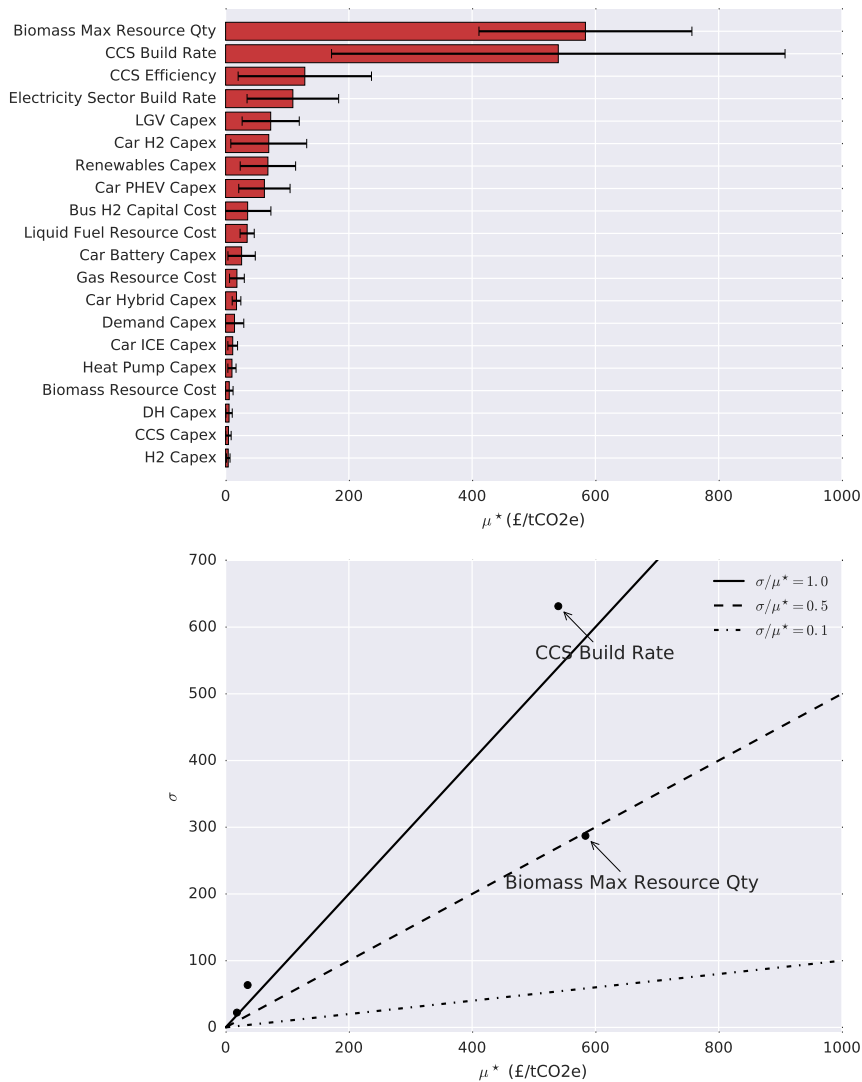


Figure 6: Influence on carbon price in 2050 of 31 inputs groups of parameters

the marginal price of CO₂ as the level of economy-wide taxation that would be required to meet the equivalent emission constraint. In Fig-

ure 6, domestic biomass availability and CCS build rate dominate the variation in CO₂ price with a μ^* four times larger than the next most influential input. In the case of the biomass and CCS constraints, it is likely that as the constraints increase in stringency (less biomass, fewer GW of CCS installed per year), the marginal alternative technology chain is very expensive indeed. The negative sign of μ (not μ^*) for the biomass resource and CCS constraints confirms this. With the exception of Liquid Fuel Resource Cost, the fossil resources do not feature highly on this plot, although the number of transport technology inputs which feature indicate this sector is particularly difficult to decarbonise.

Figure 7 shows the influence of the different input groups upon electricity production (TWh) in 2050. The size of the electricity system is an indicator of structural changes in the energy system. Previous studies (Strachan et al., 2008d; Usher et al., 2010a) have shown how electrification of heat and transport have important synergistic properties for cost effective decarbonisation of the whole energy system in the UK, albeit using a different ESOM, UK-MARKAL. Again the build rate of CCS and availability of domestic biomass are the most important for the amount of electricity production in 2050. The next most influential input groups are the Electricity Sector Build Rate and Renewables Capex — the capital cost of a group of renewable energy technologies. Interestingly, transport technology input groups also feature, indicating that there are some significant trade-offs between hybrid, hydrogen and conventional vehicles with respect to electricity. Heating technology inputs are not particularly influential. For example Heat Pump Capex has a μ^* value of 1.75 TWh.

Figure 8 shows the influence of the input groups upon a narrowly defined output metric - nuclear capacity in 2050. Only three parameters have any influence upon this output, Electricity Sector Build Rate, CCS Build Rate and Renewables Capex, with the first dominating the otherwise small values of μ^* . Nuclear capacity appears in almost all the simulations run in the ESME model indicating either that the model is under-constrained, or that nuclear presents a viable low-carbon alternative to renewables and the incumbent fossil fuel generation capacity.

Figure 9 shows the influence of the now-familiar input groups upon a derived metric from the energy system model - the proportion of electricity generation capacity that are renewable technologies, expressed as a percentage. The results here are similar to those from Figure 7, again dealing with the electricity sector. The most influential variables are Biomass Max Resource Qty and Renewables Capex followed by Electricity Sector Build Rate and CCS Build Rate. Evidently electricity generation from CCS technologies displace renewable technologies (indeed this is observed in Figure 43 on page 164). It is also interesting that parameters related to demand tech-

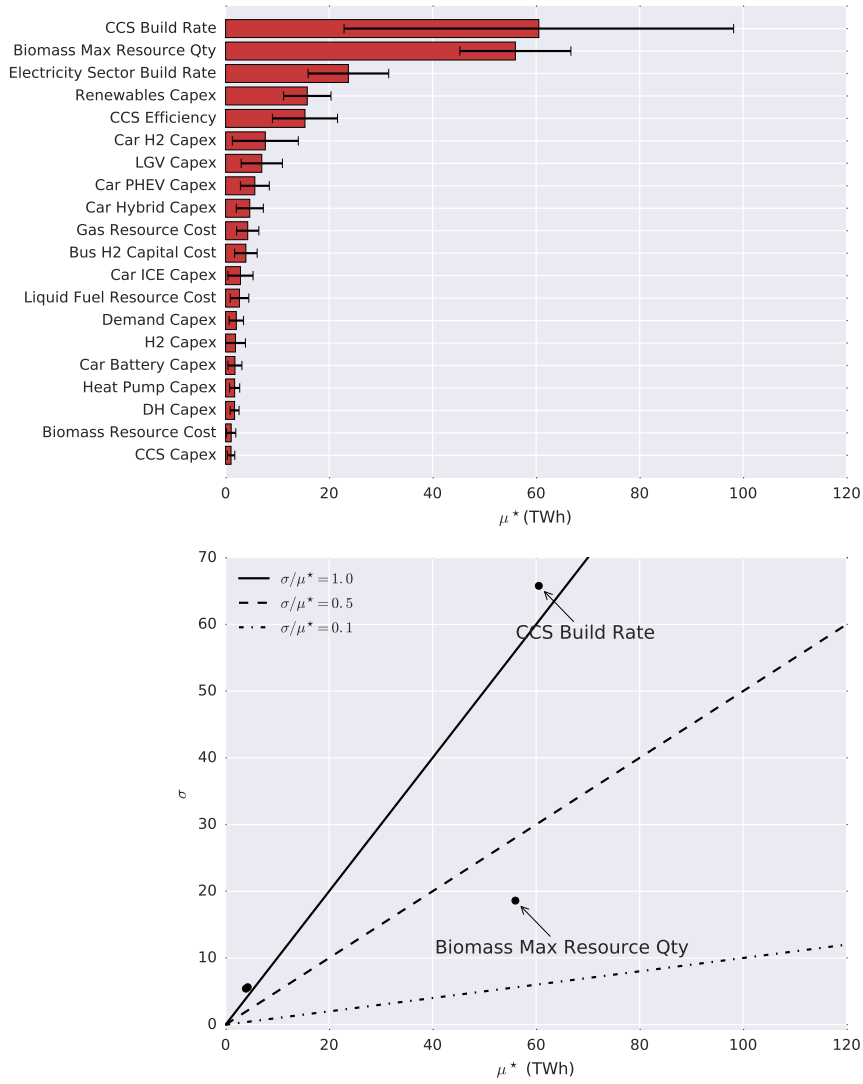


Figure 7: Influence on electricity production of 31 inputs groups of parameters

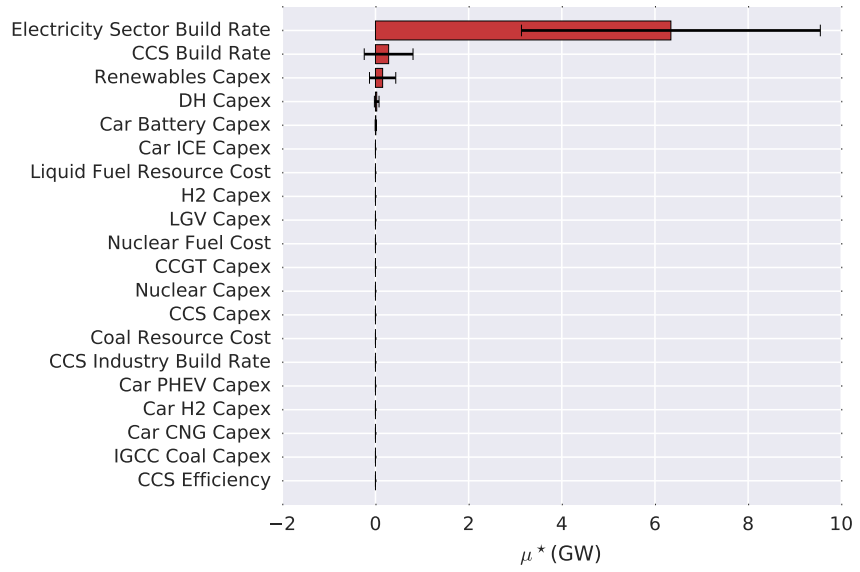


Figure 8: Influence on nuclear capacity of 31 inputs groups of parameters

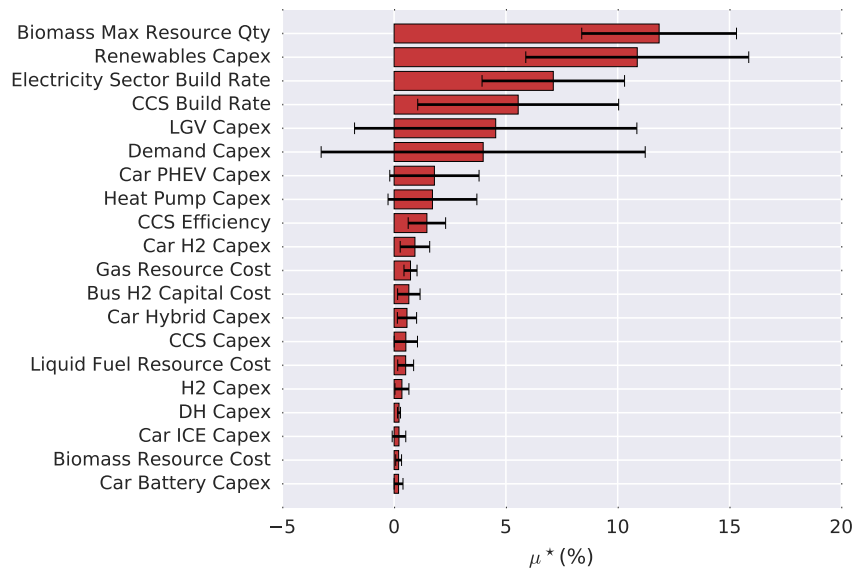


Figure 9: Influence on proportion of electricity capacity that is renewable in 2050 of 31 inputs groups of parameters

nologies, such as heat pumps, cars, and light bulbs (Demand Capex) also influence the percentage of renewable capacity in the electricity system.

4.2 RESULTS FROM THE MORRIS SCREENING: SECTORAL INVESTIGATION

The input parameters can be divided into several categories, each of which are explored further in the sections below:

- constraints on resources
- technology costs
- resource costs
- constraints on build rates
- other parameters

4.2.1 *Emission Constraints*

An initial study (see Appendix F) shows that varying a CO₂ emission constraint between a 0% and 80% reduction below 1990 levels is responsible for less variation in total cost than the cost of Liquid Fuel, and that this parameter is a significant causal driver of structural change of the modelled energy system, a fact that is not immediately obvious from the results shown in Figure 51 on page 218.

The magnitude of the constraint upon biomass resource is an important source of variability in the total cost of the energy system. In the ESME model, biomass availability enables the use of so-called 'negative emission' technologies such as IGCC Biomass with CCS. These greatly increase the amount by which the electricity system can decarbonise, thereby allowing increased emissions in sectors in which it is more expensive to mitigate.

When the constraint on carbon dioxide emissions is removed from the set of uncertain inputs, and fixed to a value within its range (in this case, the statutory reduction of 80% below 1990 levels), a corresponding reduction in output variation is viewed (compare Figures 51 and 52 in Appendix F). As a consequence of the removal of this influential variable, the rank order of the parameters changes, and the magnitude of the mean elementary effect (μ^*) for the most influential input parameters decreases. Strikingly, the cost of Liquid Fuel remains the most influential parameter, again indicating that the model is either over-constrained, or has little opportunity to substitute liquid fuel for alternative resources, or move to alternative technology chains from further downstream. An example of the latter could involve investment in alternative fuels for road passenger

Table 8: Ranking of technology cost input groups

Technology	Elec. Prod	CO ₂ price	Total Cost
Renewables	4	7	6
Car H ₂	6	6	13
LGV	7	5	7
Car PHEV	8	8	11
Car Hybrid	9	13	17
Bus H ₂	11	9	20
Car ICE	12	15	10
Demand	14	14	14
H ₂	15	20	>20
Car Battery	16	11	>20
Heat Pump	17	16	15
DH	18	18	16
CCS	20	19	19
Nuclear	>20	>20	8

transport, such as hydrogen fuel cell passenger cars, together with the upstream supply chain. This is perhaps surprising, as the extra cost of mitigating carbon dioxide emissions should, if cost effective, result in a reduction of the cost differential between fossil and alternative fuels.

The influence of the availability of biomass (Biomass Max Resource Qty) increases dramatically when the CO₂ constraint is fixed. Biomass is evidently only used under such a deep decarbonisation target, as the value of μ^* increases from £27bn to £145bn. Observing the values for μ confirms the direction of effect of the biomass constraint as negative for total cost and carbon price. Large values of σ indicate that there are significant interaction or non-linear effects associated with this parameter.

4.2.2 Technology Costs

The changing rankings of μ^* of the model outputs to the range of technology costs are shown in Table 8.

That the cost of vehicles enter the top of the sensitivity table is indicative of the lack of options for decarbonising the transport sector within the model, and the high cost differential between alternative transport systems, which would require systemic changes e.g. a wholesale move to hydrogen or battery charging networks. Under low and high carbon dioxide constraints, the demand for driving

is constant as this model version does not include demand elasticities (see Section 4.3 for results when this assumption is relaxed). Of the electricity generating technologies, the capital cost of nuclear is the most important input parameter for technology cost, reflecting the high penetration of nuclear in the model under all carbon dioxide constraint levels. The cost varies over a range of between £2100/GW and £4300/GW by 2050. However, for both electricity production and CO₂ price output metrics, the cost of nuclear does not appear in the top 20 most influential parameters. This could reflect the lack of changes between model runs — the installed capacity of nuclear is almost static across simulations indicating that there is an additional constraint, such as a build-rate constraint suppressing even more capacity. This is the case, with a maximum build rate of 1GW/year between 2020 and 2029, rising to 2GW/year from 2030 onwards. Across all the metrics, renewable technology costs emerge as the most important technologies, followed by LGV and Car PHEV.

Table 8 gives some interesting insights into the effect of battery car cost upon the model results. The cost has some importance for electricity production and helps shape the cost curve for the marginal price of carbon, but has little effect upon the total cost of the energy system. This is another demonstration of the substitution of technologies within the model results - battery cars are only viable when cheap enough, they will increase demand for electricity, but will reduce consumption of liquid fuel by displacing conventional or hybrid vehicles. Thus the net effect is not evident on the total cost of the energy system.

Additionally, the position of Car H₂ within the table shows that the use of hydrogen (H₂) has important upstream (i.e. structural) significance. If the costs of the demand technology which use hydrogen fall, then there is a multiplying effect upstream, such as using coal or biomass CCS to produce hydrogen with a net reduction or negative production in carbon dioxide.

4.2.3 Resource costs

The parameter to which the model is most sensitive is the price of liquid fuel (oil and related products), which is consumed in technologies that represent road transport vehicles, such as cars, buses, vans (LGVs) and heavy-goods vehicles (HGVs).

The model is also sensitive to the cost of natural gas, as under all CO₂ constraint levels, natural gas plays an important role in electricity generation and residential heating sectors. For example, in high carbon scenarios, the gas can be burned unabated to generate electricity as well as in gas boilers to supply thermal heat and hot water in domestic, commercial and industrial sectors. In low carbon scenarios, CCGT plants with CCS can be used to burn natural gas while

Table 9: Ranking of resource cost input groups

Technology	Elec. Prod	CO ₂ price	Total Cost
Liquid Fuel	13	10	1
Gas	10	12	4
Biomass	19	17	9
Coal	>20	>20	
Nuclear	>20	>20	18

generating a fraction of the emissions. This assumes that CCS is a successful technology and the other uncertainties associated with transmitting and storing the carbon dioxide are resolved positively. The dominance of natural gas over liquid fuels in the electricity sector is apparent from the switching positions of the two in Table 9 across the output metrics.

In unconstrained scenarios (with no cap on carbon dioxide), coal resource costs are also an important source of variation in the total system cost (see Figures 51 and 51 in Appendix F). This is due to the existing stock of coal fired power stations, the cost of coal is discounted less than fuel stocks used later in the model horizon, and that coal is one of the preferred fuels for electricity production in scenarios where the CO₂ mitigation constraint is not particularly severe. Also, due to discounting, near-term costs weigh more heavily upon the sensitivity analysis results than variation later in the time-horizon. Thus near-term variation which is locked-in could dominate the results. The grouping of technologies has largely avoided this, for example with renewables and incumbent technologies in different groupings.

The values for μ and σ show that the interaction/non-linear effects are small, indicating that liquid fuel costs and natural gas costs are treated as input costs, with few opportunities for substitution. In other words, the model uses these resources under almost all circumstances. This could be because there is no substitute for these resources, because the model is over-constrained, or because there is no cost-effective substitute, even over all combinations of alternatives.

The cost of the biomass resource is far less influential than the ready availability of the resource, while the changing cost of nuclear fuel has a larger influence upon the total cost of the energy system, than on the amount of electricity produced or the carbon price. In the former case, as a key resource for decarbonisation, the structural benefit to the system of being available is more important than the cost of the resource. In the latter case, the structural benefit to the sys-

Table 10: Ranking of max build rate input groups

Technology	Elec. Prod	CO ₂ price	Total Cost
Electricity	3	4	5
CCS	1	2	3
CCS industry	>20	>20	>20

tem of nuclear capacity is more important than the cost of generating electricity¹.

4.2.4 Build rate constraints

Build rate constraints limit the capacity of a technology that can be installed in each five-year period. This represents the uncertain availability of engineering skills, public acceptability, and the equipment and infrastructure to deliver large engineering projects to time and cost in a particular scenario. Table 10 shows the ranking of μ^* values for each of the output metrics. With the exception of the constraint upon domestic biomass availability, the build rate constraints dominate all the other parameters and have wide implications across the model output metrics. but they can act only on an individual technology. The CCS Build Rate input group contains the parameters shown in Table 33 in Appendix D. The parameter CCSStations is a constraint group in the ESME model and represents the right-hand side of a constraint which limits the yearly sum of installed capacity of all CCS plant. The maximum build rate of individual CCS plants are contained under the Electricity Sector Build Rate.

4.2.5 Other parameters

Table 11 shows how the ranking for the parameter CCS Efficiency² changes across the output results. What is particularly interesting is the CCS Efficiency is so much more influential that CCS Capex, across all of the outputs. For example, the cost of the CCS technologies has a μ^* of £3.0bn while the efficiency of the CCS technologies has a μ^* of £13.5bn. Evidently, over the range of CCS costs, the range of improvements in the efficiency of the CCS plant is more important at a system level. Note that the efficiency referred to here is the efficiency at which the plant generates electricity. This is an interesting example of the systemic trade-offs of such novel technologies

¹ Note that because the Max Build Rate parameter for the Nuclear technology is bundled within the Electricity Sector Build Rate group, it is not possible to differentiate between the importance of the nuclear capacity versus the other technologies within this group

² This is the efficiency with which the plant generates electricity

Table 11: Ranking of efficiency input groups

Efficiency	Elec. Prod	CO ₂ price	Total Cost
CCS	5	3	12

against the alternative mitigation options. Due to the unique ability of the CCS plants to sequester carbon, and particularly when using biomass as a fuel to obtain negative emissions, the efficiency will increase the already considerable benefits over and above the existing plant. Given that CCS plants are already economically viable under an 80% emission reduction, and are installed across the system, the increase in efficiency helps magnify the rate at which emissions are sequestered.

4.3 INCLUDING ELASTIC DEMANDS

The following results are from a sensitivity analysis on an elastic demand version of the ETI-ESME model. Building upon the work in Pye et al., (2014) in which the authors conduct a local probabilistic uncertainty analysis of just the elastic demand values, the results in this section are from [global sensitivity analysis](#) where the elastic demands are varied together with other parameters in a full exploration of the model input space. Due to the severe computational demands of the elastic demand formulation, the results were run for just the time period 2050. While this does mean that the full inter-temporal effects of the energy system are not captured, the results do capture the magnitude of the effects of demand reduction in the model year where the most drastic action is required to meet the steep carbon emission reduction target. The results are formed from 920 separate model runs, for the 29 elasticity values, and the 17 groups of most important variables from the results presented in Appendix F. Due to the number of groups including just one parameter, values for σ and μ were also generated giving insights into the interaction/non-linear effects for some of the variables.

Figure 10 gives the results for the reference (or baseline) scenario used to generate the base prices, from which the demand curves are subsequently extrapolated for the elastic demand mode. In this scenario, no carbon constraint is imposed, and so unsurprisingly, the cost of liquid fuel, natural gas and coal dominate the variation in total cost in for the year 2050. As the model is run for just 2050, the magnitude of the discounted total system cost is much smaller than for the earlier pathway runs, however the proportion of variation as a percentage is comparable with the results in the previous section.

Figure 11 gives the results for the total cost of the energy system under the carbon policy. Note that liquid fuel resource cost is less in-

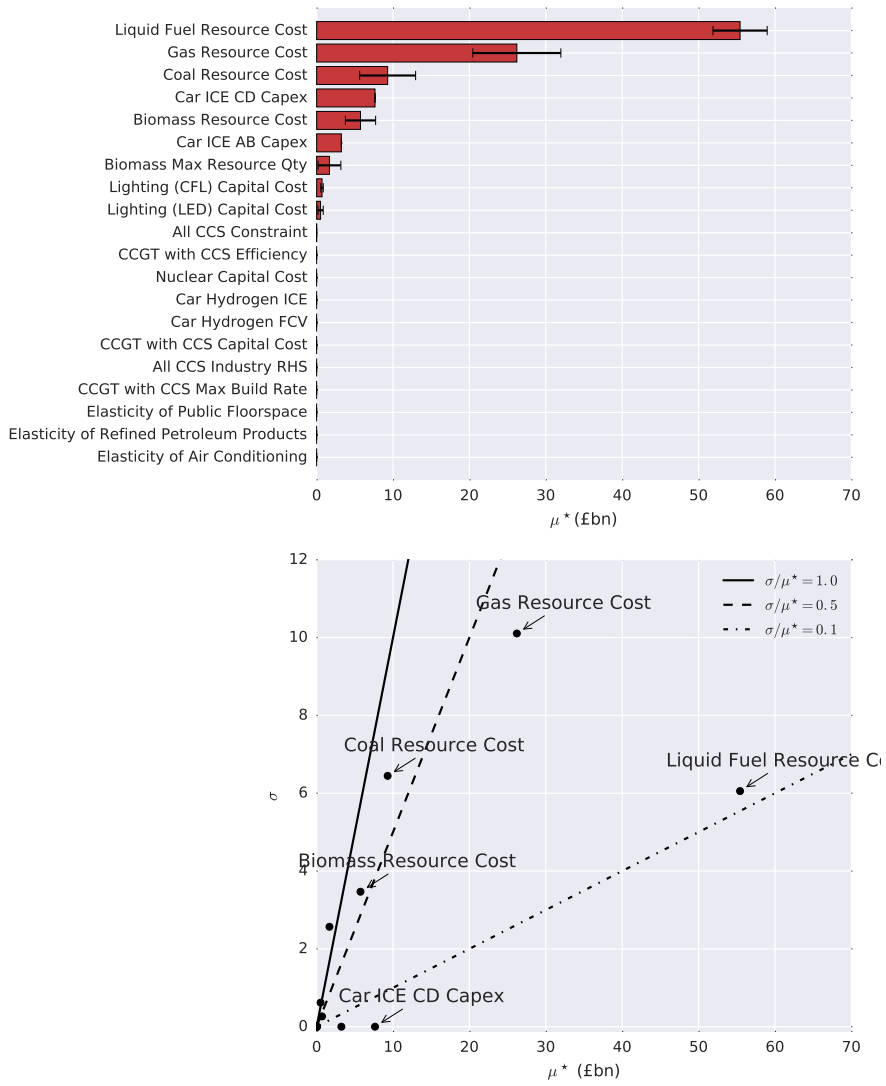


Figure 10: Sensitivity of total cost (in the reference scenario without elastic demands and with no emissions constraint)

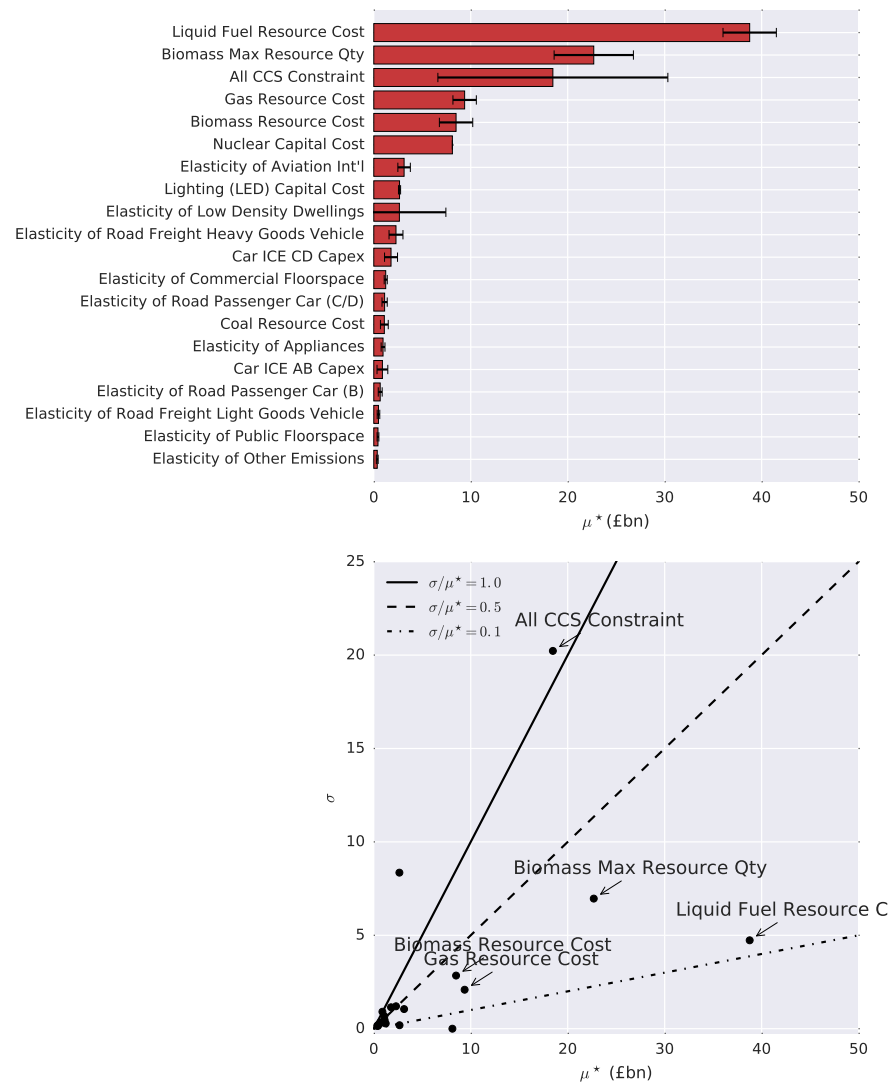


Figure 11: Sensitivity of total cost (under the policy scenario with elastic demands)

fluent under the policy case under the reference case shown in Figure 10. The imposition of the emission constraint engenders a move away from fossil fuels, so the fluctuation of liquid fuels accounts for a smaller amount of variation in the total cost than in the reference case. However, as in the non-elastic demand results in Appendix F, the biomass constraint jumps up the rankings.

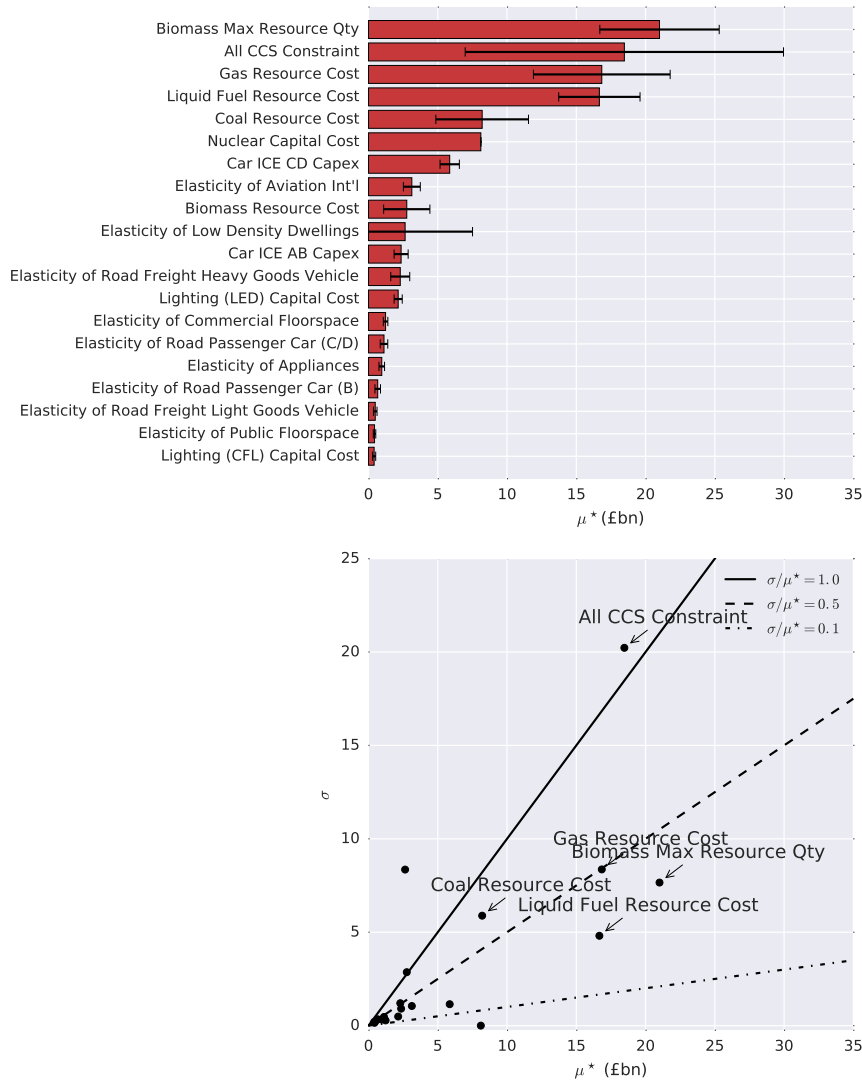


Figure 12: Sensitivity of change in consumer/producer surplus (under the policy scenario with elastic demands)

Figure 12 gives the results for the consumer/producer surplus. This result metric is computed from the difference between the total energy system cost in the reference case, and the sum of total energy system cost and change in consumer welfare due to demand changes under the policy scenario. In this figure, we therefore see the influence of each parameter upon the variation in consumer/producer surplus of the emission constraint, including both the change in the

supply and demand side parameters. In essence, this Figure shows the difference between Figures 10 and 11.

Note that the influence of the liquid fuel resource cost parameter is far lower than under either of the total cost figures, although still in the top four parameters, (Figs 10 & 11), as the variation occurs across all scenarios. Nuclear has a very small (zero) confidence interval, as the installed capacity was identical for each level of nuclear capital cost. This means that while the capital cost has an important influence on the total system cost, the effect is entirely linear, and that there is no variation around this value (nor interaction effects). Note this was revealed through *ex-post* analysis of the results, rather than interpretation of the confidence interval.

The change in consumer/producer surplus is probably a more robust result metric than the absolute values of total cost as it is relative to the changes between the scenarios, rather than due to absolute assumptions. Indeed, the variation in welfare is largely down to key decarbonisation parameters, such as the maximum available biomass quantity, and the availability of CCS technologies. The supply side accounts for a significantly larger proportion of the variation in welfare than the elasticities assigned to the energy service demands.

The results indicate that individually, no one elasticity has significant influence upon the model results. However, nine elasticities, mainly in the transport sector are influential for change in consumer/producer surplus and twelve are in the top-20 for 2050 carbon price (shown in Figure 13). When taken in sum, these elasticities are likely to appear in the top five of important parameters. In terms of magnitude, the parameters identified from the results in the previous section still dominate, with fossil resource costs and the constraints on biomass and CCS making up the majority of the variation in total cost and change in consumer/producer surplus. The reference level of energy service demands (i.e. the absolute level of energy service demand, around which demand increases or decreases) are held constant in this analysis, and it would be particularly interesting to investigate the influence of these parameters on the output metrics. This is left for future work.

One methodological issue with this approach is the content of the scenarios used to generate the reference prices from which the demand curves are generated. Typically, the reference scenario is run as a 'policy free' scenario, so without an emissions constraint or other policies, such as a renewable portfolio standard. The resulting welfare change and the magnitude of the demand responses occur as the imposition of the policies cause an increase (normally) in the shadow price of energy services, thus provoking a reduction in demand to an equilibrium price. The loss in welfare caused by the demand reduction is thus a function of the content of the reference scenario, as well as the policy changes. However, under a simulation approach,

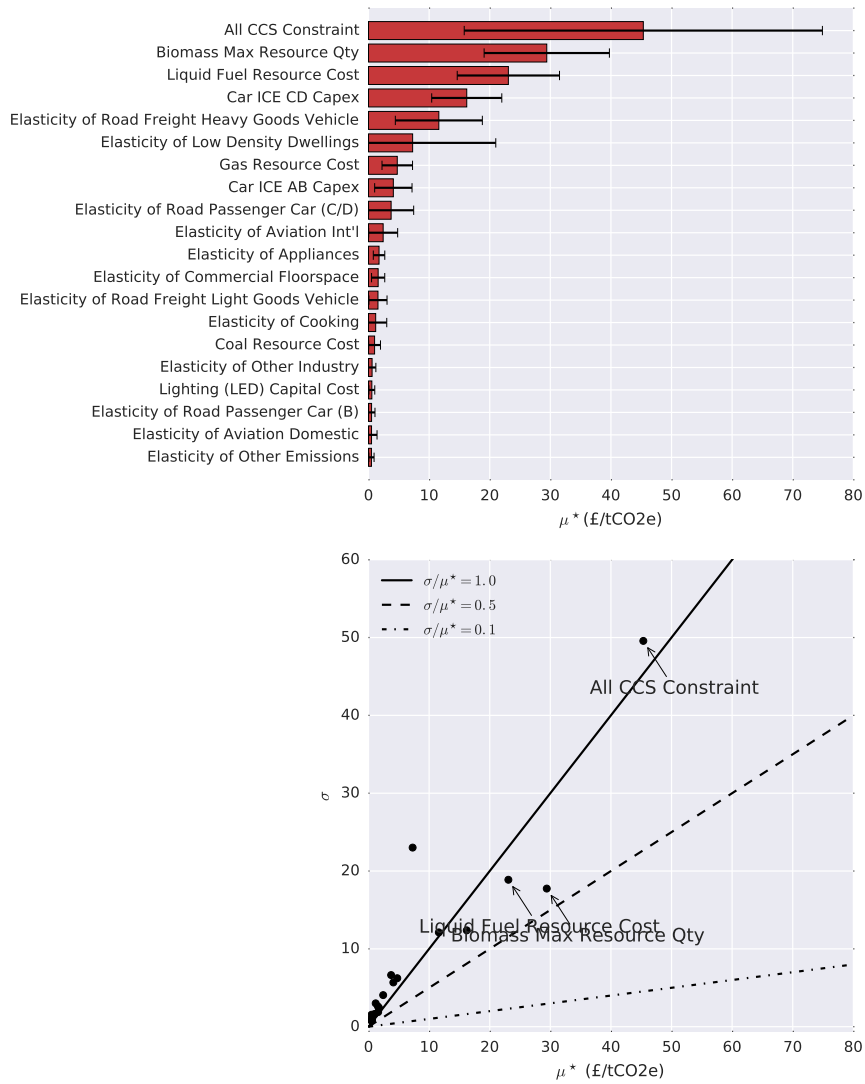


Figure 13: Sensitivity of carbon price in 2050 (under the policy scenario with elastic demands)

elements of the underlying reference scenario also change between scenarios. Thus a question arises as to whether changes in uncertain inputs, such as the cost of a technology or price of a resource should be considered as a part of the reference scenario, or a policy. The effect of the former, as suggested in Pye et al., (2014), could be that the amount of demand reduction is unchanged, as the base prices would already take into account the change in input parameters. The effect on the demand reduction of the policy would be little different to that of the original reference case. This issue has also been discussed in Strachan, (2010), who notes that

“...the inclusion of existing policies in modelling long-term decarbonisation pathways appears to be comparable to a major exogenous modelling assumption — that of global fossil fuel prices.”

The computational burden of running the elastic demand version of the model is large, particularly because for each simulation in the sensitivity analysis sample, the model must be run twice; once to generate new marginal prices for the demands, and a second time to implement the policies and compute the demand reductions. Unfortunately, this made running the model in full pathway mode with multiple key-years computationally infeasible for this thesis. However, the similarities between the results in this and the previous section indicate that the findings are consistent.

4.4 DISCUSSION

Of the most influential parameters in the ETI-ESME model, the majority of variance in the output metrics is explained by changes in just four parameters. These are Liquid Fuel and Natural Gas resource costs, the availability of domestic biomass and the build rate of CCS plant. Interestingly, the sum of μ^* for all the technology costs is significantly less than for these top three parameters. There is less inter-parameter variation in the influence of technology costs than between technology costs and the other the most important metrics. The most important technology costs are for low-carbon (renewable and nuclear) technologies and transport technologies. This indicates that after the critical system wide parameters of biomass, CCS and fossil resource prices, it is low-carbon technologies which are most important. In the context of the modelled 80% CO₂ reduction target, this makes logical sense.

The recognition of a parameter's importance is dependent upon the output metric used. This is demonstrated clearly when comparing the rankings of the input parameters over the total cost, CO₂ price and electricity production metrics. Most noticeably, while the price of liquid fuel and natural gas dominated the results for the total

cost of the energy system, they were far less influential on the price of carbon and electricity production. Clearly, the careful selection of the output metric used will influence the results from the sensitivity analysis, and the interpretation of the results is contingent on the output metric. For the purpose of this thesis, the use of total cost as the main output metric is relevant because it is this output which is subsequently used as an input to the learning model introduced in Chapter 5.

The [global sensitivity analysis](#) covered the inputs within the technical framework of the [ETI-ESME](#) model, that it was possible to vary given the available resource. These included technology costs, efficiencies, resource costs, build rate constraints and resource constraints. Notable exclusions from the [global sensitivity analysis](#) include the absolute value of energy service demands, which are otherwise varied over discrete levels in high-level scenarios, technology specific hurdle rates, the global discount rate, and any aspect of the spatial dimension (the [ESME](#) model considers 17 sub-regions of the UK). Of these, the global discount rate and the absolute demand levels are likely to be the most influential, although it is very difficult to predict how their importance will rank against the parameters which have been included in the analysis.

The findings from the sensitivity analysis pose an important question for those embarking upon a scenario analysis which exclude changes in fossil resource costs. Many scenarios conducted in past have focused on the cost of technologies, the absolute level of a CO₂ constraint, with some attention paid towards the cost of resources. In the absence of a formal sensitivity analysis approach, researchers using energy system models must use their intuition and experience to determine which parameters are important. But by quantifying and ranking the influence of the parameters using global sensitivity analyses, it is considerably easier to identify which parameters should be given preference in a study based upon scenario analysis.

One reason for the apparent sensitivity of the [ESME](#) model to the cost of the fossil resources is that the model does not include piecewise linear supply-cost curves for resources. The supply-cost curve is a horizontal line with no change in cost as a function of quantity demanded. Another reason, is that due to discounting (the discount rate is set at 3.5%), the near-term consumption of fossil fuels under all scenarios dominates the total cost result metric, while the relatively distant investment in alternative sources do not appear as prominently. However, the results show that even under a range of emission constraints, near-term oil consumption remains at a high level due to the lock-in of infrastructure, particularly in the transport sector. This result highlights the major role that oil and gas play in the current energy system, and the contribution of the large downside (and upside) risks associated with such a large consumption.

4.4.1 *Comparing Global Sensitivity Analysis to the Alternatives*

I now compare sensitivity analysis with

- Analysis of multiple scenarios
- Monte Carlo sampling
- Exploratory modelling

4.4.1.1 *Analysis of Multiple Scenarios*

The established practice of comparing multiple scenarios to manage uncertainty is widespread in the use of energy system models such as the [ESME](#) model. Such a technique does not convey to the user of the model data the complete behaviour of the model, precisely due to the fact that only very few points of the input space are covered by the approach (it is a local approach). The definition of diagnostic scenarios in Kriegler et al., (2015), while a welcome advance in the field of model comparison studies also falls short of a systematic analysis, avoiding using even a design-of-experiments style approach. In comparison, a Morris approach has significantly greater coverage of the input space, and provides a direct means by which a user of the model results can assess the most important inputs to the model, yet provides this data in a concise 'need-to-know' basis, in the form of a ranked list of input parameters. Painter et al., (2007) find that deterministic scenarios can miss important interaction effects which are identified by a sensitivity analysis. It is difficult to identify a situation in which performing a scenario analysis alone is ever preferable to one in which sensitivity analysis has been used to identify and prioritise inputs. While models are increasing in size, complexity and computational demands, most energy system models are of a manageable computational size. At the very least, a fractional factorial approach can give some initial insight into the model's behaviour at little computational cost.

4.4.1.2 *Monte Carlo Sampling*

A Monte Carlo approach generates a distribution of results given a joint distribution of input parameters. Often, quantifying input distributions is difficult, and quantifying the joint distribution of inputs, i. e. the correlations between input parameters, is also extremely challenging. This is particularly so, given the sheer quantity of data required by [ESOMs](#). The Morris sensitivity analysis approach suffers from similar issues, making the strong assumption of independence between variables. However, the advantage of the Morris approach is that significantly fewer model runs are required to obtain a good estimate of the importance of the individual input parameters, whereas

a Monte Carlo approach can require a huge number of runs to cover the input space effectively. While a latin-hypercube approach can give efficient coverage, the number of runs is still far in excess of those required by the optimised-Morris approach outlined in this thesis. Morris gives an interesting perspective on the model results as it takes into account the behaviour the model over a huge range of inputs, while the sensitivity measure μ^* , gives a direct measure of the importance of input assumptions upon an output of choice (often total-system cost, but perhaps system capacity, or a measure of system diversity). If a Monte Carlo analysis has already been performed, then the approach of Plischke et al., (2013) may be used to compute sensitivity indices from the data already generated.

4.4.1.3 *Exploratory Modelling*

Exploratory modelling is an approach which makes very few assumptions regarding the nature and form of both parametric and structural uncertainties in a model. This approach to policy modelling outlined in Lempert, (2002) and Lempert et al., (2003) requires an exploration of the input space of sample of input value combinations, running a model many thousands of times and mining the resulting data. Existing knowledge (i.e. subjective beliefs in the form of probabilities) may be incorporated, but normally over an ensemble of scenarios generated by a model, rather than for the inputs to the model. In terms of computation, this is an inefficient way to operate, as scenarios must first be generated, and those that do not fit discarded. Lempert's approach is extremely general, however, in that it accommodates multiple model structural assumptions - through parameterising the relationship between variables, or by using many different models to generate scenarios, against which signals for robustness are identified. Exploratory modelling begins to assess structural uncertainties in the model formulation as well as parametric uncertainty. Sensitivity analysis can also be configured to address structural uncertainty, through parameterising alternative model formulations, and including this parameterisation as an input to the sensitivity analysis.

One can argue that Energy system modelling sits in a middle ground, where some uncertainties can be defined using quantitative methods, but there are important borders with aspects of deep uncertainty, ambiguity and Knightian uncertainty. While many of the attributes of the energy-system planning 'problem' may fall within the domain of deep uncertainty, such an approach may be overly-conservative for making decisions in the energy sector at a national level. Chapter 5 provides some guidance in those cases where policy makers are able to take action proactively to reduce uncertainty. Sensitivity analysis is one method for identifying important uncertainties, ranking and prioritising them and then identifying optimal hedging strategies for incorporating learning into a wider investment programme.

4.4.2 *A Policy Perspective*

Another perspective on the results of the [global sensitivity analysis](#) is to establish the implications for policy. There are a number of policy options applicable to the parameters to which the [ESME](#) model is most sensitive. The objective function of the [ESME](#) model is formulated so as to minimise the total cost of the energy system, an objective which is intended to correspond to the preferences of policy makers to transition to a future energy system in as cost-effective way as possible.

Hughes et al., (2012) suggest that there is a distinction between uncertainties that exist due to the inaction of decision makers and those that exist due to decision makers' inability to control. The level of CO₂ emissions is an example of the former - given a lack of political will, an uncertain emissions constraint is a significant factor in the variation of the total system cost. The level of CO₂ constraint provides significant uncertainty to not only the total cost, but also the structure of the energy system and the portfolio of technologies required. However, setting a CO₂ constraint via a regulation, as outlined in the UK's Climate Change Act (2008) therefore reduces the variation in total cost, and thus the investment risk to which society is exposed. This is subject to a host of caveats pertinent to the 'real-world' including: enforcement of the emissions cap, gaming or societal belief that the emissions cap are indeed enforceable.

On the other hand, policy makers within the UK have relatively little control over the prices of oil and gas resources. The UK is a price taker for such fossil fuel commodities, especially given dwindling production rates for oil and natural gas from the North Sea. Furthermore, the markets for oil and natural gas are global (or at least regional) in nature, and thus the UK is subject to the same fluctuation in oil prices as the rest of the world or region. The natural gas market is less fluid in nature, although with the recent increase in liquidised natural gas (LNG) terminals, global trade in natural gas is increasing. However, while policy makers have little control over the price of resources at their entrance to the country, taxes and subsidies can be used to alter the prices seen by consumers as well as providing a source of revenue (in the case of taxes). The results from the sensitivity analysis indicate that the prices of oil and gas are key determinants of the total cost of the energy system under all levels of CO₂ constraint. However, the modelling results do not suggest that oil and gas resources ever reach prices high enough to encourage a wholesale substitution that is optimal across the whole system. This is perhaps surprising, given the opportunities to transform the transport system, the primary consumer of liquid fuels, through a transition to hydrogen or battery-electric vehicles. Both of these technology options are well-detailed in the [ETI-ESME](#) model. Ultimately, only

detailed analysis of the specific scenarios in which high oil and gas prices occur over combinations of other parameters (such as cheaper battery-electric vehicles or hydrogen technologies) would shed light on this. The Morris screening and other sensitivity analysis methods can help lead the analyst to these conclusions, that could easily be missed given the number of inputs that can potentially influence the model results.

The sensitivity results also show that the response of the cost of the energy system to the cost of liquid fuel and natural gas are largely linear - again reinforcing that there is little, other than reducing demand or accelerating a switch away from fossil fuels that can mitigate the large influence of these parameters. This also highlights a potential failing of the [ESME](#) model, in that there are few mitigation options available which provide a credible mid-term alternative to oil and natural gas. If this is indeed reflective of the real UK energy system, then the implications are that a low oil price would be in the national interest regardless of the imperative to move to a low-carbon energy system.

4.4.3 *Learning about Uncertainty*

Relatively few dynamic uncertainties appear in the top rankings of the sensitivity analysis of the ETI-ESME model. Those that do include costs of renewable technologies (grouped), CCS Build Rate, Biomass Max Resource Quantity and Electricity Sector Build Rate. For the renewable technologies, reducing the capital cost, or improving the efficiency has the similar effect of altering the relative balance between the technologies for their total costs (the discounted sum of capital, operation and maintenance and variable costs). This reduction in total system cost is achieved through either the direct reduction in total costs because either the technology is cheaper, or an increased investment in the capacity of that technology thus substituting other more expensive technologies, or a combination of both.

In the case of CCS technologies, it seems that the availability and efficiency of the technology is more important than the cost when using electricity production and carbon price output metrics. Evidently the benefits to the energy system as a whole outweigh the otherwise expensive costs when in direct comparison to alternative technologies. As defined in the model, the build rate constraints are binding, and any efficiency improvements of the technology gives an added bonus. It is the structural benefits and the enabling factor of the CCS technology which provides its main value to the energy system. So perhaps there is a role for learning in resolving the uncertainties surrounding the deployment of CCS - as represented by the build rate constraint in the [ESME](#) model?

The availability of domestic biomass is a critically important uncertainty in the [ESME](#) model. In the sensitivity analysis, the parameter is moved over a wide range of resource availability, from 0 TWh/year forcing expensive imports to be used, to a very high 600 TWh/year, which is the upper limit of plausibility in the existing literature. The results show that over all the output metrics considered, domestic biomass is of considerable importance to the [ESME](#) model. Biomass can be used across the energy system, from the production of bio-fuels, co-firing with coal in CCS plants, production of H₂ via biomass gasification with CCS and the firing of biomass in district heating systems. This systemic flexibility is likely to contribute to the importance of biomass availability, as constraints on biomass limit the possible structures of the energy system. Resolving uncertainty around the availability of biomass early would allow early investments that facilitate rather than hinder an energy system in which biomass plays a strong role.

This systemic structural importance is common to both CCS and biomass. Constraints on either of these are important both on an absolute level, but also through time, due to the path dependency of large infrastructure. It will be interesting to investigate how the dynamic availability of both CCS technologies and biomass play out over time, and how these two technologies interact. The timing of the availability of biomass and CCS was not investigated as part of the sensitivity analysis, but could well be included in future iterations.

4.5 SUMMARY

In this chapter, I have presented the results of a sensitivity analysis of the [ESME](#) model. Grouping 119 inputs into 31 groups, I generated sensitivity metrics using just 320 simulation runs of the model. I have shown how the technique can be used to provide useful supporting insights into how the model behaves. By comparing sensitivity metrics derived from different model outputs, an informative meta-study on the model is possible, and the insights link with observations and the experience of other model results presented in the literature.

One caveat of the results is that the coverage of model inputs is incomplete. However, while the continued addition of inputs to existing groups would not result in an increase in the number of simulations needed to produce the sensitivity analysis metrics, the results indicate that the majority of inputs produce little or no impact upon the outputs. Thus the probability of each additional input materially affecting the results decreases as the number of inputs included increases. In fact, only four of the input groups, which cover just five of the 119 model parameters, account for the majority of the variation in model outputs. Of these inputs, two are resource costs, one is a constraint on groups of key decarbonisation technologies, and one

is a resource which is a key contribution to one configuration of a low-carbon energy system.

The next two chapters build upon these findings, by integrating the Build Rate of CCS and the Availability of Domestic Biomass into a learning model linked to the [ETI-ESME](#) model. The results from the sensitivity analysis enable the study to be focused where a reduction in uncertainty can have the most impact, confident in the fact that concentrating on a select few parameters will not miss important insights. Both the biomass and CCS uncertainties have similar properties. Both are influential uncertainties, and are strongly non-linear or interacting with other uncertain inputs, (distinguishing between the two is not possible with the method of Morris), CCS particularly so. However, while the costs of liquid fuel and natural gas are strongly influential, their effect is less important for decarbonisation of the energy system, and the majority of their effect is linear/non-interacting in nature.

LEARNING UNDER DYNAMIC UNCERTAINTY: METHODOLOGY

In this chapter I outline the formulation and description of the learning model. The aim of the learning model is to quantify the value of resolving the critically important uncertainties for the UK energy system that were identified by the sensitivity analysis in Chapter 4.

In Section 5.1, I discuss the modelling of [research projects](#) — the predominant device by which learning can take place. Section 5.2 contains the mathematical formulation of the learning model. And in Section 5.3 I demonstrate the behaviour of the learning model using simple examples. These examples nonetheless show some of the behaviours we hope to find from the model. In Section 6.1, I define the [research projects](#) used in the thesis, building upon the results from the sensitivity analysis. In the following Chapter 6, I run the full learning model [ESME](#) linkage using the two critical uncertainties identified in Chapter 4.

The steps to quantitatively evaluate uncertainties are as follows. Discrete (e.g. low, central, high) values, taken from the upper and lower bounds and median values in the [ETI-ESME](#) data are used for each of the critical uncertainties. A [revenue function](#) is constructed using costs from running the [ESME](#) model multiple times, holding the value of the uncertain parameters at different levels at different points in time. Then, the learning and [ESME](#) models are linked through the [revenue function](#) to find the optimal timing and level of the associated R&D projects.

Figure 14 summarises the model linkage. The learning model is linked to the [ETI-ESME](#) model in two ways. The successful result of individual [stages](#) of a [research project](#) are linked to the discrete values of the critical uncertainties. In the diagram, a [three-stage research project](#) is associated with a decrease in cost. The successful result of each [stage](#) results in a predefined decrease in cost. The learning model decides if and when to conduct a [research stage](#). The choice to embark on the [stage](#) of a [research project](#) is governed by the expected net benefit of the resolution of the corresponding uncertainty in the [ETI-ESME](#) model.

5.1 LINKING RESEARCH PROJECTS AND DYNAMIC UNCERTAINTIES

Before detailing the formulation of the learning model in Section 5.2, I first explain how [research projects](#) are modelled in this thesis, how the

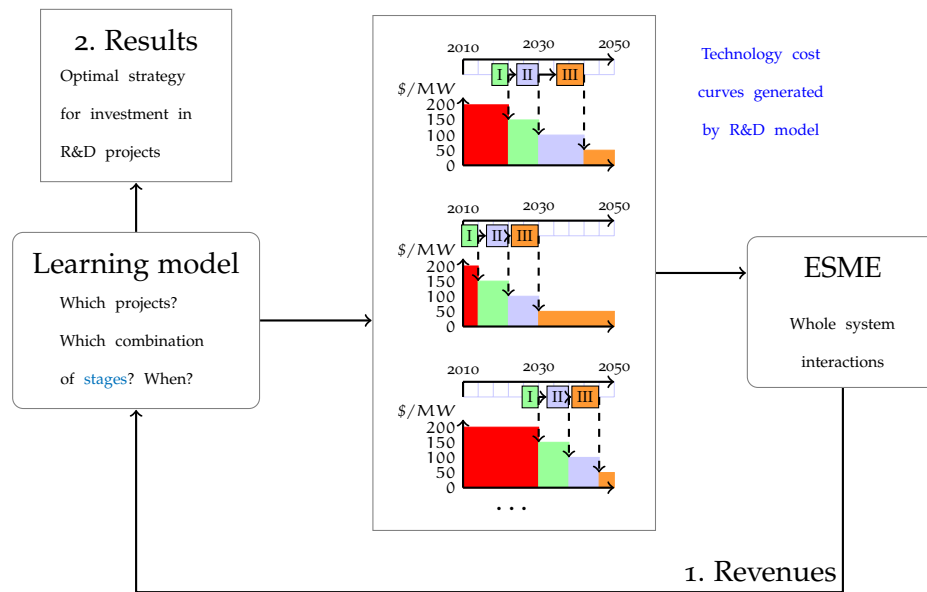


Figure 14: A schematic showing the linkages between the ETI-ESME model and the learning model

decision dependent uncertainties affect the structure of the [scenario tree](#), and how the use of sequential and consecutive [research projects](#) simplifies the approach.

5.1.1 Modelling Research Projects

In Chapter 4, two critically important parameters were identified; the availability of domestic biomass and the build rate of CCS. Both of these uncertain parameters are sufficiently broad as to merit an equivalent approach. In this thesis, the device for evaluating learning about these uncertain parameters is to assume that there is a *research project* which, if successful, could resolve the uncertainty associated with each parameter. The precise nature of the research and details that would comprise such a project are not investigated, as the aim of this approach is to estimate the financial value of such an endeavour rather than specify the precise nature of the [research projects](#) that would lead to the desired outcomes. Essentially, this approach links an investment — the cost of the [research project](#), with an uncertain result — the success or failure of the project, for each of the critically uncertain parameters.

By modelling the learning process as a research project, an existing modelling approach, outlined in Section 5.2, can be used. The use of a [research project](#) as a device for learning also raises the potential for investigating the division of a project into a series of sequential [stages](#), later [stages](#) only being started upon the successful completion of earlier [stages](#). This raises prospects of a real option style approach.

Table 12: Typology of Uncertainties for Decision Dependent problems

Exogenous		
Decisions Affect	No Link	
Example	Demands, Prices, Market	
Structure of Tree	Known & fixed	
Endogenous	Type I	Type II
Decisions Affect	Probability Distributions	Timing of Observation
Example	Measure for longer	Decide to measure
Structure of Tree	Depends on decisions	

In addition, the approach could investigate a portfolio of competing projects.

Table 12 compares the definition of endogenous and exogenous uncertainty drawing upon the definition of Goel et al., (2006).

First, it is important to establish how the chosen modelling approach is able to correctly represent the research process. A [research project](#) may be divided into one or more consecutive [stages](#), each of which can either succeed or fail. The probability of the success or failure of an individual [stage](#) is unaffected by the decision over whether to start the [stage](#), and is specified exogenously. However, the timing of the observations of the result of the [research project stages](#) are dependent upon the decision to start the particular [stage](#) of the [research project](#). The nature of the uncertainties in multi-stage [research projects](#) are clearly endogenous and decision-dependent. The endogenous process is the link between the decision to invest in a [stage](#) of a [research project](#) and the observation of whether that [stage](#) succeeds or fails — known as Type II [endogenous uncertainties](#).

Note that the alternative approach, modelling Type I [endogenous uncertainties](#), where the probability of a [stage](#) is affected by investment, requires an alternative set of assumptions. These assumptions concern the relationship between investment and the probability of success of a project.

5.1.2 The Structure of Research Projects

The purpose of this section is to try to understand the underlying mathematical logic of the selection of [research projects](#) as they are defined in the learning model. We define a [research project](#) as a sequence of binary [stages](#), each of which can succeed or fail. This can be represented using a binary tree, where each node represents a decision to continue with the project, or to halt. Figure 15 shows a decision tree for a two-stage [research project](#). The probabilities of the project succeeding are signified by p and q . X through Z signify

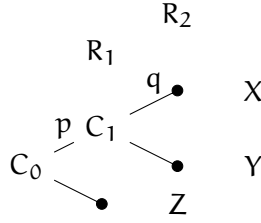


Figure 15: A tree that represents a two-stage research project

the results, C represents the cost of the stage and R the revenues of the stage. The research project succeeds at X , fails after starting the second stage at Y , and fails after the first stage at Z . The net present value:

$$\text{NPV} = pR_1 - C_0 + pqR_2 - pC_1 \quad (19)$$

In Section 5.4.1, I show the optimal strategy of front-loading uncertainty when faced with two otherwise identical research projects. In that example, $R_1 = 0$, $R_2 = 10$, $C_0 = C_1 = 1$, $p = 0.1$ and $q = 1.0$ for Project 2 with p and q swapped for Project 1. Substituting the values from that example into the equation above gives $\text{Project}_1 = 1.0 \cdot 0 - 1 + 1.0 \cdot 0.1 \cdot 10 - 1.0 \cdot 1 = -1.0$ $\text{Project}_2 = 0.1 \cdot 0 - 1 + 0.1 \cdot 1.0 \cdot 10 - 0.1 \cdot 1 = -0.1$ Project 1 is less than Project 2. The difference between the two NPVs is a result of the effect upon the cost of the second stage. If the uncertainty is back-loaded, then you are more likely to have to pay C_1 even if the project fails overall. When $p = 1.0$ instead of 0.1, then the project is equivalent to a one-stage project with sequential costs, whereas when $p = 0.1$, the NPV is computed over two stages with the opportunity to abandon the project after the first stage.

If in the two-stage research project shown in Figure 15, the revenue of a research project is only recouped after successfully completing the entire project¹, i.e. $R_1 = 0$, then the expected net present value can be computed as follows:

$$\text{PV}_{\text{flex}} = pqR_2 - pC_1 - C_0 \quad (20)$$

This valuation accepts that the project can be abandoned after conducting the first stage of the project, hence the flex subscript.

If an interim revenue R_1 is received after the successful completion of the first stage, the net-present value NPV is computed as in Equation 19. Note that as the revenue from the first stage R_1 increases or exceeds the cost of the second stage C_1 , the terms cancel out, and the net-present value of the project matches that of a single-stage project.

¹ This is likened later to a “big pharma” distribution of research revenues, in contrast to interim or “partial revenues” which accrue throughout the research stages

However, if the project cannot be abandoned, the NPV is computed differently:

$$PV_{fix} = pqR_2 - C_1 - C_0 \quad (21)$$

In other words, the costs of conducting the project are viewed as one aggregate sunk cost and the separate **stages** do not exist for purposes of evaluation. In this case, for the project to be selected, the probability weighted revenue needs to counterbalance this sunk cost, thus ensuring $NPV > 0$.

The difference between Equations 20 and 21 gives the value of the flexibility to abandon a project:

$$V_{flex} = PV_{flex} - PV_{fix} \quad (22)$$

$$= C_1 - pC_1 \quad (23)$$

As the probability p is bounded as follows, $0 < p \leq 1$ then $PV_{flex} \geq PV_{fix}$ and $C_1 \geq V_{flex} > 0$. So V_{flex} is largest when p is very small and C_1 is very big, indicating that two-**stage research projects** where the majority of uncertainty is front-loaded, will be most beneficial. To maximise the value of a flexible two-**stage** project, a cheap initial **stage** with low-chance of success would be selected to be conducted first, delaying a more certain, but expensive second **stage**. This mirrors what we see in the real-world: the use of cheap prototype projects, or proof-of-concept, before committing larger resources to a project.

If we do the same for a three-**stage** problem again with all revenues only recouped upon successful completion of the final **stage**:

$$PV_{flex}^{3stg} = pqrR_3 - pqC_2 - pC_1 - C_0 \quad (24)$$

$$PV_{fix}^{3stg} = pqrR_3 - C_2 - C_1 - C_0 \quad (25)$$

$$V_{flex}^{3stg} = PV_{flex}^{3stg} - PV_{fix}^{3stg} \quad (26)$$

$$= C_1 + C_2 - pC_1 - pqC_2 \quad (27)$$

$$= C_1(1 - p) + C_2(1 - pq) \quad (28)$$

$$= C_1 + C_2 - p(C_1 - qC_2) \quad (29)$$

we see that the value of flexibility is the product of the probability of failing each **stage** and the avoided cost of conducting that **stage** (Equation 28). Because the probability of failing rises as the number of **stages** increases, e.g. because $(1 - pq) > (1 - p)$, the cost of later **stages** will increasingly influence the value of flexibility if the costs of conducting each **stage** are similar e.g. $C_1 \approx C_2$. An alternative way of interpreting this is to say that to maximise the value of flexibility, the mostly costly **stages** should be delayed, regardless of the associated revenue.

Assuming an infinitely divisible project, in which as many homogeneous **stages** of equal cost and probability of success can be defined as required; is it intrinsically better to have many rather than fewer

stages? If a research project is assumed to have p likelihood of success, with a cost of C , then the project can be divided into T stages. Each of the stages would have a probability $\hat{p} = p^{\frac{1}{T}}$ of successful completion, with the overall likelihood of success being equal to p . Likewise, each stage would cost $\hat{C} = C/T$. The value of flexibility would then be computed using the following:

$$PV_{flex} = \hat{p}^T R_T - \hat{C} \left[\frac{1 - \hat{p}^T}{1 - \hat{p}} \right] \quad (30)$$

$$PV_{fix} = \hat{p}^T R_T - \hat{C} T \quad (31)$$

$$V_{flex} = PV_{flex} - PV_{fix} \quad (32)$$

$$= \hat{C} T - \hat{C} \left[\frac{1 - p}{1 - p^{\frac{1}{T}}} \right] \quad (33)$$

$$= C - \frac{C}{T} \left[\frac{1 - p}{1 - p^{\frac{1}{T}}} \right] \quad (34)$$

Due to the increasing exponent as the number of stages increase, the costs of the latter stages become increasingly close to one another, thus there are rapidly diminishing returns to adding extra stages. For example, if $T = 2$, $C = 10$ and $p = 0.1$, $V_{flex} = 3.4$ ($\hat{p} = 0.31$). Further increasing T to 3 results in a small increase of V_{flex} to 4.4 ($\hat{p} = 0.46$). As T increases, V_{flex} tends towards a maximum of $C - C/T$.

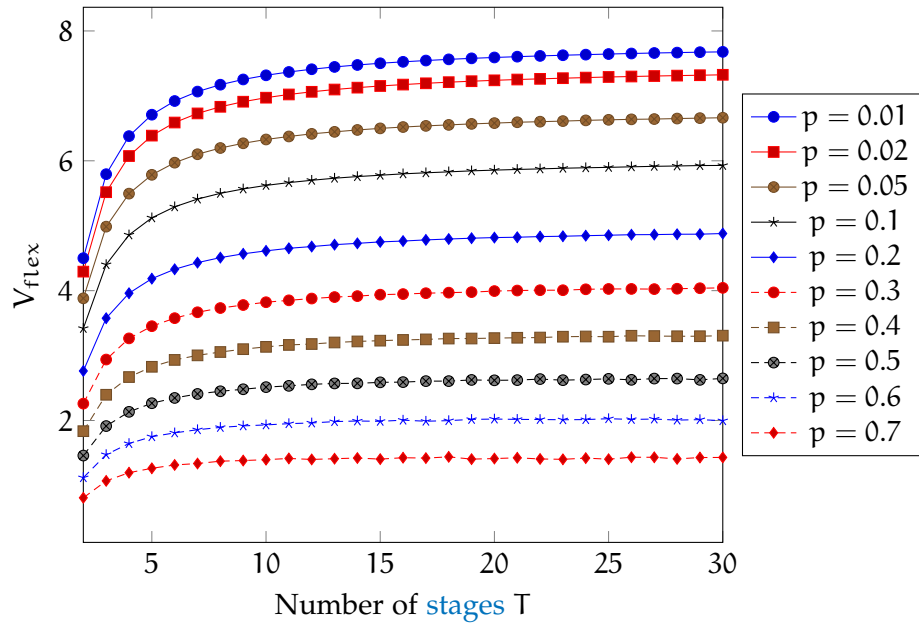


Figure 16: The value of flexibility increases with the number of stages T over a range of probabilities p of research project success. The cost of completing the research project $C = 10$.

Figure 16 shows how the value of flexibility increases with the number of stages T over a range of probabilities p of research success. In

cases with very low probabilities of success, there is a large increase in the value of flexibility as the number of [stages](#) is increased. In cases where the probability is greater (e.g. 50%), then the marginal increase in V_{flex} is much less. Note that these values do not include the increase in transaction costs that would be associated with evaluating many stages in a project.

5.1.3 *Endogenous Uncertainty and Scenario Tree Structure*

Stochastic programming is a modelling approach which incorporates [exogenous uncertainties](#) into the objective function. A fixed [scenario tree](#) is defined which represents the joint distribution of uncertain parameters, in each time period. This is relatively easily formulated into a tractable [deterministic equivalent](#) formulation, which can then be solved using existing commercial solver software.

However, under Type II [endogenous uncertainty](#), the timing of the observation of the result of the [stage](#) of a [research project](#) changes as a consequence of the decision over if and when to start the [stage](#). This is problematic for traditional ways of formulating a stochastic programming problem, as the uncertainty is no longer exogenous, the [scenario tree](#) is no longer fixed; the probabilities in each time period are dependent upon the value of the decision variables.

One option for modelling endogenous uncertainty via stochastic programming is to define multiple [scenario trees](#), only one of which is chosen depending upon the value of a decision variable. However, this problem quickly explodes in size and in many cases, limiting the techniques that can be used to solve the optimisation problem (Dupačova, 2006). An alternative is to generate a [scenario tree](#) and switch on-and-off the [non-anticipativity constraints \(NACs\)](#) which define the timing of uncertainty resolution across scenarios in the [deterministic equivalent](#) of the model. The [NACs](#) are linked to the values of decision variables. This is the approach taken by Colvin et al., (2008) and the approach I adopt in this thesis. The cost is that only relatively small problems are solvable even adopting branch-and-bound solution techniques. For example, Colvin et al., (2010) solve problems of seven R&D projects each of six [stages](#).

Colvin et al., (2008) outline two ways in which their technique reduces the size of the problem. The first is to model binary results (e.g. success/failure) of [research project stages](#). The second is to enforce an order in which decisions must be made, as this reduces the total number of possible results if, for example, a decision is only taken when the previous decision ends in success.

In the case where decisions are non-sequential (e.g. the assumption of sequential project [stages](#) is relaxed) the [scenario tree](#) cannot be pruned as the probability of a project's success is not conditional upon a previous project. The projects can be conducted in parallel

and in any sequence. However, results can only be defined as success or failure. For example, Figure 17 shows that there are four outcomes with two project, when projects are non-sequential. If the first project fails (node 1) it is still possible to conduct the second project and either succeed (node 4) or again fail (node 3).

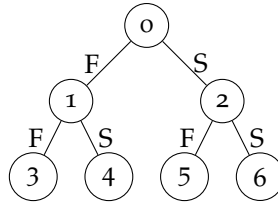


Figure 17: Scenario tree: Binary results, non-sequential projects

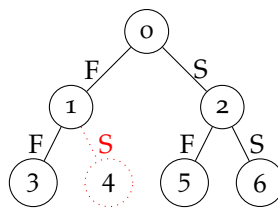


Figure 18: Scenario tree: Binary results, sequential projects

Instead, consider the decision to invest in one or both of two sequential R&D projects. If the first R&D project fails, then the second project cannot succeed, and therefore would not be started. If the first project succeeds, then there is still a chance that the second project will fail. Figure 18 shows how failure at the first project removes the option of the second project. This reduces the number of outcomes from 4 to 3. The reduction in number of outcomes is more pronounced as the number of sequential *stages* increases. For example, Colvin et al., (2008) use a mixed integer programming formulation to represent a multi-*stage* stochastic programme for the selection of three-*stage research projects* in pharmaceutical research, where the results of successive trials are irrelevant if a previous trial fails. This then enables the decision tree that would be of size $8^{|T|}$ to be pruned to $4^{|T|}$, where T is the number of trials.

Note that despite the apparent limitations of discrete (binary) parameters, and consecutive *research project stages*, this approach does allow modelling of continuous parameters (not just binary) as shown in Figure 19. This allows gradual or partial successes of *research projects* to be modelled. For example, a partial cost-reduction of a technology with an associated likelihood. In this thesis, partial successes are modelled, where the results of an initial *stage* reveal the partial value of an uncertain variable, and the option to continue reveals the full benefit of the uncertain parameter.

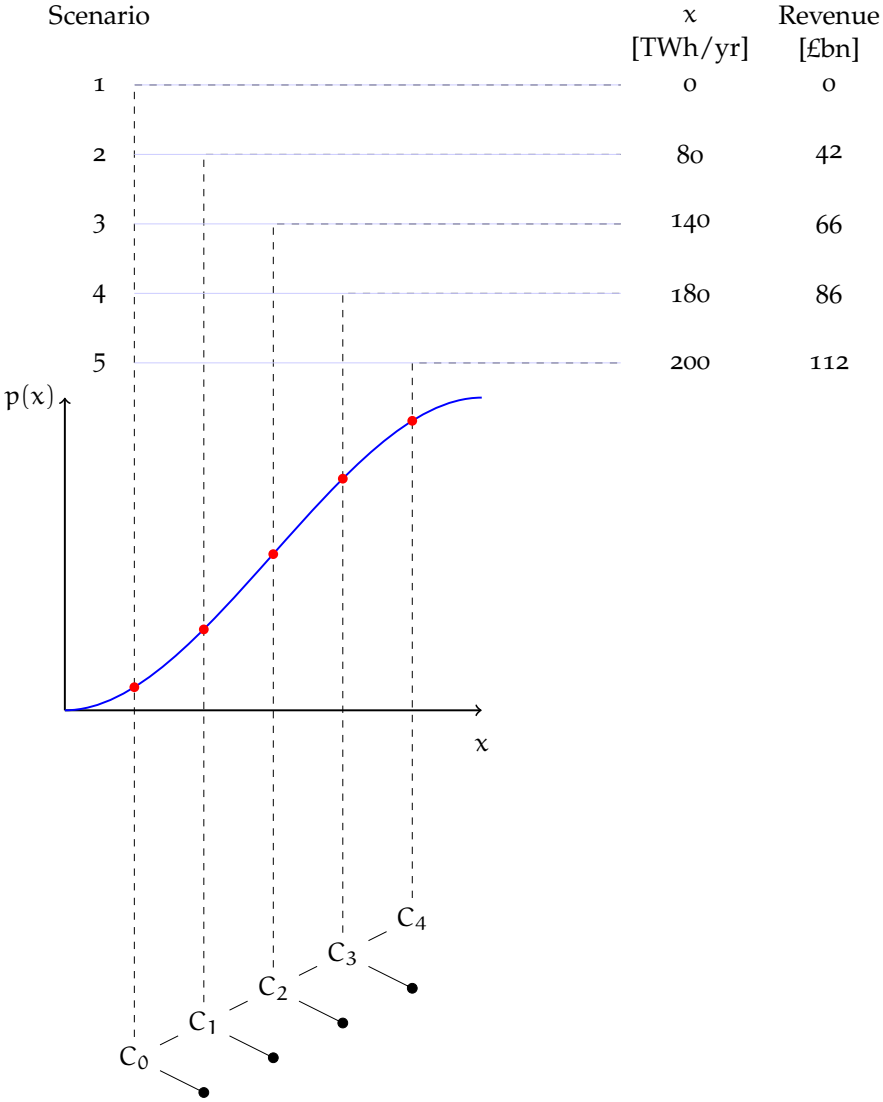


Figure 19: A decision tree that represents a four stage research project, with equivalent (continuous) cumulative probability distribution. The true probability distribution would be discrete.

5.2 THE LEARNING MODEL

In this section, I describe the learning model, and give the mathematical formulation. In the following sections I refer to **STAGES**, **OUTCOMES** and **SCENARIOS**.

STAGES refer to the individual consecutive phases of a **research project**. A cost and probability is associated with each stage of a research project.

OUTCOMES refer to the enumeration of the results of the **research project stages** in the learning model (pass/fail) *independent of when or whether investment in the project occurs*. The **outcomes** are associated with probabilities, derived from the probability of success of each **stage** in a project (see Section 5.2.1). The **outcomes** are equivalent to the leaves of a binary tree as shown in Figure 18.

SCENARIOS refer to the combinations of successful research projects and **stages** for each key year in the **ESME** model. The scenarios define the steps that make up the **revenue function** from the **ESME** model (outlined in 5.2.2) and quantify the value of each successful research investment strategy.

The model is formulated as a multi-stage mixed-integer stochastic programme, where uncertain outcomes of the **research projects** are represented using a multi-stage **scenario tree**. In this model, the **scenario tree** takes a special structure, which reflects the assumption of consecutive **research projects**, and the novel structure of the conditional **non-anticipativity constraints (NACs)**, which enable the modelling of dynamic uncertainty. The decision variables identify a multi-stage hedging strategy through the combination of **research project stages** and time periods, in which an optimal recourse action is taken in response to the **outcome** in each time period. The research model horizon is resolved in five-year periods, similar to the **ESME** model. A number of **ESME scenarios** need to be run to define the **revenue function** for the learning model. This is a function of the number of ‘key years’ t^{KY} a subset of the **ESME** time periods. This allows the decision space of the learning model to be kept manageable, at the expense of a loss of granularity. In the results that follow in Chapter 6, key years are defined around the middle of the time horizon — between 2020 and 2040 — as this is where a number of the critical decisions concerning the structure of the energy system take place (see Usher et al., 2012, for example).

The main decision unit within the model is the **research project**. A **research project** is described using the following input parameters:

Cost_{ij}^R the cost of conducting the **stage** j of project i

τ_{ij} the duration of the **stage** j of project i

\hat{p}_{ij} the probability of the **stage** j of project i completing successfully

h'_z the revenue received from successful completion of the [stage](#) according to scenario z^2

Each [research project](#) may consist of one or more consecutive [stages](#). Each [stage](#) j within a project i has a duration τ_{ij} , a cost Cost_{ij}^R and a probability of succeeding \hat{p}_{ij} . Importantly, the successful completion of a [research project](#) is governed by the successful completion of the consecutive and ordered [stages](#) which comprise the project. Taken together, the probability of successfully completing each possible combination of [stages](#), defines the exact probability distribution of the problem.

The objective function is formulated as a maximisation of the present value of the research pipeline. This means that financial constraints are not taken into account — there is no research budget — if a research programme is profitable, given the expected revenue minus costs, and there are non-financial resources available to do so, the [research project](#) is undertaken.

The formulation of the learning model taken from Colvin et al., (2008, 2010) is outlined in Appendix B. The adjustments to the original formulation, particularly of the research project revenues through a link to the [ESME](#) model are outlined below. Additions to the formulation are listed in the following sections. In:

§5.2.1 I describe the formulation for generating the outcomes for the combinations of [research projects](#), [stages](#) of the projects, and the number of targets (uncertain parameters) of the [research project](#), and in

§5.2.2 — 5.2.3 I describe the linking procedure which connects the input data for the R&D project with the scenarios run in the ETI-ESME energy system model and the subsequent generation of a revenue function governing the value of successfully completing a research project, or combination of research projects, in each key year

5.2.1 Outcome Generation

This section outlines the procedures for composing the [scenario tree](#) which represents the exogenously defined uncertain outcomes of the [research projects](#) and their component [stages](#).

² In the paper of Colvin et al., (2008), a revenue is associated directly with the completion of the final stage of the project. In the formulation shown later, the revenue is a function of the results from the [ESME](#) model, and differs according to the timing, but also the combination of successful stages across one or more projects. This combination of project stage is encapsulated in the scenarios. Thus the inter-dependencies between research project outcomes are modelled through the [ESME](#) model.

The number of outcomes $|S|$ is a function of the number of **research projects** $|I|$ and the number of project **stages** $|J|$:

$$|S| = (|J| + 1)^{|I|} \quad (35)$$

To enumerate the outcomes, we first compute the **stride** l_i from the index of the result, for each project i :

$$l_i = \begin{cases} 1 & \forall i = 1 \\ l_{i-1} \cdot |\text{OC}| & \forall i > 1 \end{cases} \quad (36)$$

The result oc of project i in **outcome** s is:

$$oc_{si} = \left(\left\lfloor \frac{s-1}{l_i} \right\rfloor \bmod (|\Omega_i|) \right) + 1 \quad (37)$$

Research projects only continue if the previous **stage** was successful, so the number of possible events can be described by a discrete random variable of sample space Ω_i e. g. {I-F, II-F, III-F, III-P} when there are three **stages**, that is the set of possible results Ω_i for project i . I-F is short-hand for “project failed after the first **stage**”, II-F is shorthand for “project succeeded after the first **stage**, but failed after the second **stage**.”

To compute the **outcome** probabilities p_s associated with the **research project** results oc , the probabilities of each **stage** \hat{p}_{ij} must be accumulated through the **scenario tree** as follows:

$$p_s = \begin{cases} \prod_{j \leq j(oc_{si})} \hat{p}_{ij} & \forall i \in oc_{si}^P \\ \prod_{j=j(oc_{si})}^J (1 - \hat{p}_{ij}) & \forall i \notin oc_{si}^P, j = I \\ \prod_{j=j(oc_{si})} (1 - \hat{p}_{ij}) \prod_{j < j(oc_{si})} (\hat{p}_{ij}) & \forall i \notin oc_{si}^P, j > I \end{cases} \quad (38)$$

5.2.2 Implementing the Changes in Parameter Values in the ESME model

The sensitivity analysis identified critically important uncertain parameters for the **ESME** model. **Research projects** in the learning uncertainty model are associated with a subset of these parameters in the **ESME** model. Upon the successful completion of a **stage** in the **research project**, the parameter value is amended to reflect the newly observed value linked with the **outcome** of the **research project**.

In **ESME**, changes in parameter values over time caused by the successful completion of **research projects** are handled by adjusting the bounds on constraints related to those parameters. The two critical parameters, are biomass availability and CCS build rate. Each of

these are altered automatically in the [ESME](#) model when building the revenue tree for the learning model. The revenue tree consists of every possible combination of the defined research projects, given the chosen key years t^{KY} .

The [ESME](#) model is constrained (see Section 5.2.4), so that the model cannot anticipate the successful completion of a [research project](#) before the information is revealed. This approximates more closely the actual value of embarking on a [research project](#), separate from investments in the energy system, while ignoring the potential information that could arise from the decision to invest in research to an energy system decision maker. Using the [ESME](#) model in perfect foresight mode, the model has perfect knowledge of the result of each [research project outcome](#), this over-estimates the revenue from this [research project](#), as the energy system is optimised prior to the time period in which the result of the [research project](#) is realised. In non-anticipativity mode, the model is unable to anticipate the result of research projects until the time period in which the information is available. This results in a more realistic pattern of investment.

As described in Section 3.4.1, in [ESME](#), changes in parameter values over time are handled by multiplier functions and scaled by the same value across all years at run-time. In simulation mode, [ESME](#) simulates only the parameter value for the final model year, thus maintaining the shape of the parameter value over time, while the final and intervening values may differ. This poses a problem for the approach used in this thesis, as it is necessary to represent discontinuous steps in the parameter values over time, rather than the smooth transitions included in the standard model data-set. There are two options. The first is to, reverse engineer the desired parameter value in each year from the initial parameter value and index value. The alternative, which is the method used in this thesis, intervene later in the model compilation process and directly edit the parameter values before solving the model.

In the initial version of the model, we keep all technology learning, due to R&D, for technologies other than those modelled in the R&D model deterministic and fixed over time at the expected value (i.e. the mean of the values) as in Santen et al., (2016). Thus the objective function is a scalar for each combination of technology parameter values in each time period. This greatly simplifies the soft-linking of the [ESME](#) model and the learning model. I leave for future work the integration of both exogenous and endogenous uncertainty into the model formulation.

The approach taken requires that the starting point of the research programme is the most heavily constrained option, with each subsequent [stage](#) reducing the total cost of the energy system. It is necessary to make an assumption as to what the unresolved value of the uncertainty should be, but in most cases, this is straightforward.

For example, in the case of biomass availability, the unresolved value is the minimum extent of the range, while the research programme defines *stages* which reveal if more biomass is available. And with CCS Build Rate, each successive research *stage* increases the upper bound of the build rate for CCS.

5.2.2.1 Computing the project, time period and cost-step combinations

For project i , and with a maximum of one *research project* taking place per time-period $t > 1$, the total number of ESME scenarios $|Z|$ that need to be performed are:

$$|Z| = \left(\sum_{j=0}^{|I|} \binom{|T|-1}{j} \right)^{|I|} \quad (39)$$

However, if the time horizon in the ESME model runs are of a coarser temporal resolution than those in the R&D model (i.e. we restrict the time periods in the ESME model to a subset of those in the R&D model), which is necessary given the computational demands of running each ESME *scenario*, multiple time-step reductions must be allowed in each time period after the first. This increases the number of scenarios that should be run in the ESME model. However, as the scenarios in the ESME model are performed over fewer key years T^{KY} , there is a net reduction in the number of ESME scenarios.

$$|Z| = \left(\sum_{j=0}^{|I|} \frac{(|T^{KY}| + j - 2)!}{j! \cdot (|T^{KY}| - 2)!} \right)^{|I|} \quad (40)$$

Note that this is a rather crude formulation and does not take into account that some combinations are not possible due to the duration of research *stages*. There is an opportunity to further reduce the number of runs of the ESME model that are required but this is left for future work.

For each of the $|Z|$ scenarios, all the combinations of *research projects*, successful *stage* results and key year are generated using a combinatorial algorithm. For example, if two years are defined, and one *research project* with only one *stage*, the scenarios would be as shown in Table 13. In this case, *scenario 1* is always null, representing the case when no *research project* takes place. In *scenario 2*, the project completes successfully in 2030. And in *scenario 3*, the project completes successfully in 2050.

Note the distinction between the actions that can be taken in the learning model, and the implications for the parameter values in the ESME model. In the example given above, the possible actions are to start the research period in any time period before either 2030 or 2050 (e.g. at least 2030 minus the duration of the research project).

Table 13: The ESME scenarios enumerate the outcomes across the key years

Scenario	1	2	3	Project 1
Outcome				
1	•			(1, F)
2	•	•	•	(1, P)
Project 1, 2030	0	1	0	
Project 1, 2050	0	1	1	

These define the three [scenarios](#) and are associated with a different pattern of investment in the [ESME](#) model if (and only if) the project is successful³. However, there are only two possible [outcomes](#) for the learning model - either the project succeeds or fails after the first (and only) [stage](#). In the event that the project fails, then only scenario one is relevant as the effect is the same for the [ESME](#) as if no project was started (the parameter values remain unchanged).

In summary, the scenarios represent an exhaustive list of the possible combinations of successful research project stages and key years. The outcomes represent an exhaustive list of the possible combinations of success/failure of the research project stages. As the scenarios also enumerate the key years, for all cases where $|T^{KY}| > 1$ there are always more scenarios than outcomes.

The next subsection describes how the mapping between [outcomes](#) and [scenarios](#) are established using constraints, and how this links the value of these [scenarios](#) to the [outcome](#) of uncertain [research projects](#).

5.2.2.2 Populating the ESME Revenue Function

After the combinations for the [ESME](#) scenarios are generated, values associated with the [research project stage](#) results are assigned to the linked [ESME](#) parameters (see Fig 14 on page 96).

The ESME model is then run for each scenario z from 1 to $|Z|$ computed in Equation 40. The resulting vector of objective function values h_z are then subtracted from the reference [scenario](#) h_1 in which no research project is undertaken, to compute the revenue received when a particular combination of research [stages](#) is completed in key year t_{KY} .

$$h'_z = h_1 - h_z \quad (41)$$

³ This assumes that the project has an effect upon a parameter value which then influences the pattern of investment in the [ESME](#) model.

This modified vector of research revenues h'_z is then used in a modified revenue Equation 42 to replace that in Equation 65 (see Appendix B):

$$\text{Rev}_s^T = \sum_z (1 - W_{s,z}) \cdot h'_z \quad (42)$$

5.2.3 Incorporating the ESME revenue function into the R&D model

After computing the vector of revenues, a constraint is needed to ensure that the correct revenue is chosen for each stochastic **outcome** in the learning model.

A binary decision variable W_{sz} is used to select the appropriate reward from the vector of revenues. The value of W_{sz} is changed to one when there is a match between the research project stage combinations held in $b_{zit^{KY}}$ and the decision variables which represent the actual pattern of investment made $\sum_j Y_{ijts} \cdot R_{sij}$.

$$(0 + M \cdot W_{sz}) \geq \sum_j Y_{ijts} \cdot R_{sij} - b_{zit^{KY}} \geq (0 - M \cdot W_{sz}) \quad (43)$$

$$\forall z, i, t^{KY} \in T^{KY}, s$$

This is a big-M constraint, where M is a large number, and must be larger than the sum of the project **stages** for each technology. If the middle term is zero, then W_{sz} is forced to zero. If the middle term is non-zero, then W_{sz} can be 1. The constraint holds for each of the key years t^{KY} . Because the decision variable Y_{ijts} (see formulation in Appendix B) is active for all years after the successful completion of an R&D project, the constraint is specified in terms of this variable.

Only one ESME revenue z must be chosen for each **outcome** s .

$$\sum_z W_{sz} \cdot d_{sz} = (|S| - 1) \quad \forall s \quad (44)$$

Not all **outcome/scenario** pairs need to appear in the constraints above, as some results are not possible in particular outcomes. We therefore build a subset of **outcome/scenario** pairs d_{sz} based upon the relationship between the possible results shown by each **scenario** z for each **outcome** s . This reduces the model size by a small amount, by restricting the number of matches that have to be made for the rewards in the variable Rev_s^T .

$$d_{sz} = oc_{si}^P \geq b_{zit^{KY}} \quad \forall s, i, z, t^{KY} \quad (45)$$

Table 14: The structure of the relationship between outcomes and scenarios, enforced by the constraint shown in Equation 45. The final two rows enumerate the results and key years associated with each of the six scenarios.

Scenario	1	2	3	4	5	6	Results
Outcome							Project 1
1	•						(1, F)
2	•	•	•				(2, F)
3	•	•	•	•	•	•	(2, P)
Project 1, 2030	0	1	0	2	1	2	
Project 1, 2050	0	1	1	2	2	2	

There are further opportunities for reducing the number of constraints in problems with multiple key years. For example, in Table 14, it is not possible to differentiate between scenarios 1 and 3 until 2050, as they are identical in 2030. On the other hand, if the value 0 (indicating that the project has not been successful, or that no project has been started) matches in 2030, we can exclude scenario 2 from the constraint set. Using a similar technique to the non-anticipativity constraints, we can define an adjacency matrix to help reduce the number of constraints, by identifying identical combinations with a key-year less than or equal to the current key year.

Namely, the constraint shown by Equation 43 that enforces matching of the decision variable X_{ijts} to the combinations of ESME scenarios $b_{zit^{KY}}$ only needs to be generated for those combinations that are unique in each key year t^{KY} as shown in Equation 46.

$$e_{zt^{KY}} = \left\{ \sum_{z'}^Z (b_{zit^{KY}} = b_{z'it^{KY}}) = 1 \quad \forall \quad i, t^{KY} \leq t^{KY}, z \right\} \quad (46)$$

5.2.4 Introducing Non-Anticipativity Constraints into the Deterministic ESME Scenarios

To ensure that the revenue function obtained from the ESME model is not an over-estimate of the value of reducing an uncertainty, it is important that the model is unable to anticipate the reduction in a parameter value. Deterministic models have the attribute of ‘perfect foresight’, such that the values of the decision variables are optimised not just for the current period but over the whole model horizon. To avoid this, the following equations ensure that the decision variables of the ESME model are held constant across specific scenarios.

The time period up to which the constraints are active are dependent upon the research decision that the [scenario](#) is computing. Specifically, the threshold of the non-anticipativity ESME scenarios follows that of the key years, and are defined by the pattern of unique combinations of projects. For each of the decision variables in the ESME model, the years prior to the key years are frozen to the values from the most similar previous [scenario](#). The links between scenarios, and the threshold year in before which the decision variables are frozen are defined by the Equations 47 and 48.

$$t_z^{my} = \arg \min t' \left(\sum_{it} b_{zit} - b_{zit'} \right) < 1 \forall z \quad (47)$$

$$z_z^{ms} = \arg \min z' (b_{zit_z^{ms}-1} - b_{z'it_z^{ms}-1}) \forall i \quad (48)$$

These constraints are implemented in the [ESME](#) scenarios by copying the values from the [scenario](#) \hat{z}_z^{ms} in [scenario](#) z for all values of $t < t_z^{my}$.

5.3 EXPLORING THE LEARNING MODEL: ONE RESEARCH PROJECT

In this section, I introduce the behaviour of the learning model through simple examples. As I explained in Section 5.2, [research projects](#) are modelled as follows. A [research project](#) comprises one or more [stages](#). A cost is incurred at the beginning of each [stage](#) and a revenue is received upon successful completion of the entire [research project](#). The [research project](#) duration determines the length of time between cost and revenue. A probability defines the likelihood of each [stage](#) succeeding. Any individual [stage](#) has only two outcomes - succeed or fail. A succeed outcome may influence a change in a parameter in the [ESME](#) model, or allow continuation to the subsequent stage (if any). A fail outcome results in the immediate halt to the [research project](#).

How to divide up a research project into individual stages is explored further in Section 5.4, as this is a complex topic deserving of significant future investigation. After compiling the sets of research projects I and dividing those projects into stages J , the learning model is parameterised with the following:

r Discount rate

h'_z discounted revenue for scenario z

Cost_{ij}^R Cost of [R&D stage](#) j for project i

\hat{p}_{ij} Probability of [R&D project stage](#) i, j succeeding

ρ_{ijk} Non-financial resource requirements

τ_{ij} Stage duration

ρ_k^{max} Resource availability

which are also listed in the nomenclature on page xxv.

In the following examples I replace the [ESME](#) model, which in the full results in Chapter 6 gives the revenue, with a very simple capacity expansion problem with two technologies. The two alternative technologies are used to illustrate the differences between the ways of structuring a [research project](#). One of the technologies is a cheap incumbent. The other technology is a currently expensive, but innovative technology, whose cost is reduced below that of the incumbent if a [research project](#) succeeds. The [revenue function](#) of this problem is the difference between the total cost of the optimal investment decision in either case; the [research project](#) succeeds or fails.

The simplest case to consider is a [research project](#) of one [stage](#) over one or more time periods. I then introduce a [research project](#) divided into multiple [stages](#). This introduces the flexibility to abandon a project without incurring the cost of the whole project in the event of an early failure. However, I first discuss the structure of [research projects](#) in an algebraic sense.

5.3.1 One Stage, Two Time-Periods

Two competing technologies, one incumbent and one innovative, are available to meet the increase in demand for new capacity from 10 GW in 2030 to 20 GW in 2050. Both technologies have an economic and technical lifetime of 20 years. A government agency may invest in a one-[stage](#) research program to reduce the cost of an innovative technology from £M200 /GW/year to £M100 /GW/year. The conventional technology costs £M120 /GW/year with 100% certainty.

The first step (mimicking the use of the [ESME](#) model) is to determine the value of the research programme. Given the two key-years, 2030 and 2050, for which the value of these investment alternatives have been computed, there are three possible combinations of capital costs for the innovative technology as shown in Table 15.

The revenue of alternatives can be computed by solving the optimisation problem to minimise total cost of the investments. Under the first result, in Table 15, investment will focus on the conventional technology as it is cheaper in both periods. Under the second result, the 10 GW of capacity in 2030 will come from conventional technology with the remaining 10 GW capacity in 2050 from the innovative technology. Under the third result, all 20 GW of capacity will come from the innovative technology. Using a 3.5% discount rate, discounting to 2010 and assuming capital costs are overnight, the total costs are as shown in the fourth column of Table 15. The difference between values in the total cost column then give the potential revenue from mak-

Table 15: Scenarios for R&D projects. Numbers in brackets are capital costs for the (sub-optimal) investment.. These scenarios correspond to those shown in Table 13.

	R&D Results	CAPEX (£M/GW/year)		Total Cost (£M)	Revenue (£M)
		2030	2050		
1	fails or no R&D	120 (200)	120 (200)	12900	0
2	succeeds in 2050	120 (200)	100 (120)	12200	700
3	succeeds in 2030	100 (120)	100 (120)	10800	2100

ing an R&D investment. Note that the revenues are not derived from the reduction in the cost of the innovative technology in isolation (£M200 /GW/year to £M100 /GW/year), but from the marginal technology, in this case, the conventional technology at £M120 /GW/year. Assuming investments are irreversible, so early retirement of the conventional technology is not possible, the innovative technology does not come into play unless the capital cost is reduced to below that of the conventional technology, when it then entirely displaces investment. This means that the success or failure of the [research project](#) wholly determines whether investment in the innovative technology goes ahead. Furthermore, without investing in the [research project](#), there is no chance that the innovative technology can be deployed. The cost of the innovative technology is a decision-dependent uncertainty.

Given this input data, the questions we can now turn to are:

- Under which conditions should the government agency invest in an R&D project?
- if so, in which period should the R&D project take place?

The answer depends upon the probabilities and costs associated with the [research projects](#).

If the success of the R&D project is certain, then a decision maker would be willing to pay a premium of up to £2100M to achieve success by 2030 and £700M to achieve success by 2050. However, as the value of the 2030 option is always higher than the 2050 option, then investment would always take place as early as possible.

As the probability of success decreases, the premium decreases proportionally. If the probability of success $p(1_s) = 0.2$ then investment in the R&D project would go ahead if the project cost were less than $p(1_s) \times \$2100M = \$420M$ if the project could be completed by 2030, or $p(1_s) \times \$2100M = \$140M$ if the project could only be completed by 2050.

5.3.2 *Two Stage, Two Time-Periods*

I now investigate a modification of the previous example. The R&D project associated with the innovative technology is now split into two *stages* instead of one, each of which bears an equal share of the cost of the *research project*, but with revenue only associated with successful completion of the final *stage*. I refer to this distribution of revenues as “big pharma”, from the paper of Colvin et al., 2008. In pharmaceutical research, the benefit of a series of trials on a drug are only realised at the end of the successful completion of the process when the drug is commercialised.

In the previous example, the only decisions were i) whether to invest in the R&D project or not, and ii) in which time period to invest.

By separating one large project into multiple small project *stages*, we now introduce a degree of flexibility, i.e. the option to halt a research programme at an intermediate *stage*. We express this as a multi-stage stochastic programme, where the decision to invest in a *stage* of a *research project* is a node on a decision tree (see Fig. 20). The tree branches according to the number of *research projects* and *stages* of each project. The multi-stage stochastic programme computes the expected net present value of investing in the portfolio of one stage-*research projects* by collapsing this probability weighted decision tree. Each *stages* in the multi-stage problem correspond to the resolution of the uncertain results of the *research projects*.

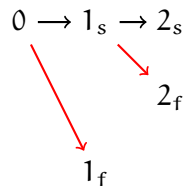


Figure 20: The decision tree associated with one research project split into two *stages*. If the first *stage* succeeds (1_s), the decision maker can continue to *stage* two, in which case the set of results is $(2_s, 2_f)$.

We can set up this problem so that it is equivalent to the previous one. If the probability of reducing the cost of the technology 2_s is also equal to 0.2, the sum of other two possibilities $p(2_f)$ and $p(1_f)$ must equal 0.8. As a reward is only associated with 2_s , these problems are equivalent in all but one aspect — the decision maker could choose to halt the *research project* after *stage* I on the realisation of either 1_s or 1_f . The introduction of multiple *stages* also has the effect of spreading the payments for the research over the project in addition to introducing the flexibility to abandon a project if it fails early.

This time, the decision maker is willing to pay upto £572M for a project which matures in 2030 or £191M for a project which matures in 2050. Because of discounting, the project which matures earlier will

always be preferred, as this will pay surplus for all project costs less than the premium in this case.

Note the increase in value of this flexible project, versus the previous project of one [stage](#). Through the division of the [research project](#) into two [stages](#), rather than one, and given the option to abandon the project, the value of conducting the project increases from £420M to £572M. This additional value reflects the value of flexibility of the two-stage project over and above the one-stage project, with everything else held equal. The exact reason for this increase in value of a project, precipitated only by its division into two stages rather than one is investigated further in Section 5.1.2 where I look into the mathematical relationships that underpin the learning model.

This finding also raises the question of determining the optimum balance of costs and probabilities for [research projects](#) greater than one [stage](#). The following section explores the preference for front- or back-loading uncertainty in a project, where all the rewards are received at the end of the project.

5.4 STRUCTURING MULTI-STAGE RESEARCH PROJECTS

It is also interesting to consider which is more desirable – a project with most uncertainty at the beginning, or the end of the research process? Consider two identical [research projects](#) of two [stages](#) each, and only one key year - the period in which the financial reward of the [research projects](#) are assessed at the end of the time horizon. Under these conditions, the number of scenarios and outcomes happen to be the same. Again, the distribution of revenues is of the “big pharma” style, only upon successful completion of the final [stage](#) of the [research project](#).

The successful completion of the second [stage](#) of either project by the end of the time horizon results in a reward of £10B. Each [stage](#) costs £1B to conduct. The probability of success of the projects and [stages](#) are shown in Table 16.

Table 16: List of research projects

Project	stage	Cost (£B)	Duration	p(Success)
Project 1	1	1.000	1	1.00
Project 1	2	1.000	1	0.10
Project 2	1	1.000	1	0.10
Project 2	2	1.000	1	1.00

The probability of each project succeeding is $1.0 \times 0.1 = 0.1$. However, the first [stage](#) of project 1 has a probability of 1.0, while the second [stage](#) has a probability of 0.1. For project 2, the first [stage](#)

has a probability of 0.1 of success, but the second **stage** is certain to succeed.

The **scenario tree** that results from these inputs is pruned because of the probabilities of 1.0 (outcomes which contain the events (project 2, **stage** 2, fail) or (project 1, **stage** 1, fail) are excluded). Of the nine possible outcomes, only four remain with probabilities 0.81, 0.09, 0.09 and 0.01 respectively.

Table 17: The structure of the relationship between outcomes and scenarios, enforced by the constraint shown in Equation 45. Only outcomes 2,3,8,9 have a probability > 0 with the inputs shown in Table 16. F=Fail, P=Pass.

Scenario	1	2	3	4	5	6	7	8	9	Outcomes	
Outcome										Project 1	Project 2
2	•	•								(2, F)	(1, F)
3	•	•	•							(2, P)	(1, F)
8	•	•		•	•		•	•		(2, F)	(2, P)
9	•	•	•	•	•	•	•	•	•	(2, P)	(2, P)
Project 1	0	1	2	0	1	2	0	1	2		
Project 2	0	0	0	1	1	1	2	2	2		

Table 17 shows the relationship between the **OUTCOMES** and the (**ESME**) **SCENARIOS**. Not all outcomes map to all the decisions which can be made within the scenarios, due to the restricted combination of time periods and projects.

I now present a series of examples of increasing complexity. The intention is to demonstrate the kinds of strategies that are likely to be displayed in the results of the learning model and to understand why these strategies are selected. Note that discounting is not taken into account in the following examples. Discounting will affect the results through the incentive to delay expenses, while bringing forward revenues. This is most likely to affect the timing of projects of multiple **stages**. The reason discounting is excluded from these analyses, is that the results tend to be rather sensitive to the discount rate chosen, blocking the appearance of some strategies.

5.4.1 Mutually Exclusive Research Projects

If the problem is to choose the optimum strategy to complete only one **research project** the projects are mutually exclusive. No extra revenue is earned if both projects are completed successfully, but costs are incurred for each research **stage** in which we invest (see Table 18 for revenues in this case). Under strict mutual exclusivity, only one project should be completed by the end of the time horizon. Embark-

ing on a hedging strategy in this case is effectively penalised, because successful completion of [stage 1](#) for one project means that a success at [stage 2](#) of the other project would result in a revenue of zero.

Table 18: The revenue is a function of the observed combination of outcomes (of which there are nine permutations) at the end of the time horizon.

Scenario	1	2	3	4	5	6	7	8	9
Project 1	0	1	2	0	1	2	0	1	2
Project 2	0	0	0	1	1	1	2	2	2
Revenues (£B)	0	0	20	0	0	0	20	0	0

The optimal strategy under these conditions is to first conduct the first [stage](#) of project Project 2, wait and see as to what is the result and then:

IF PROJECT 2 = SUCCESS Continue with Project 2, [stage 2](#)

IF PROJECT 2 = FAILURE Abandon Project 2 and start Project 1, [stage 1](#)

A visual representation is shown in Figure 21.

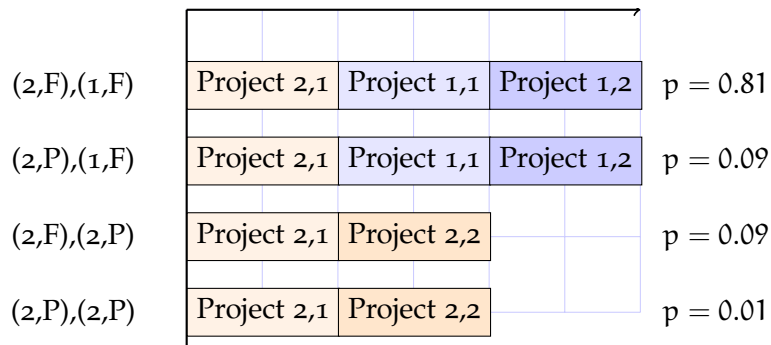


Figure 21: Optimal strategy under mutually exclusive revenue function

The reason for this can be understood from the structure of the outcomes shown in Table 17⁴. In this table, [outcome 8](#) has an equal probability of 0.09 to [outcome 3](#). However, the cost of observing the result of [outcome 8](#) requires investment in four project [stages](#), as opposed to [outcome 3](#), which requires investment in just three project [stages](#). So because it is less costly to reveal the uncertainty of Project 2 — the uncertain result is revealed in [stage 1](#), after just one [stage](#) of

⁴ Note that the values in the revenue table (e.g. Table 18) are only valid for the outcomes listed in Table 17 if the decision variables chosen mean that a project is selected. For example, scenarios 1 through 9 could all be valid under [outcome 9](#), but only scenario 1 or 2 are valid under [outcome 2](#).

investment — the optimal strategy is for investment in the Project 2 project earlier than that of project Project 1.

Because no reward is received for the **outcome** in which one project is partially successfully completed with the other project successfully completed, the optimal strategy is to invest first in the project that is most likely to fail (or for which it is cheapest to find out if it will fail) and then switch to the other project in the event of a fail, otherwise stick with the original project.

5.4.2 *Mutually Exclusive Projects - A Portfolio*

The previous either/or approach effectively penalised outcomes that included a partial success in one project together with a complete success in the other.

If partial successes are included in the **revenue function** (see Table 19), the option to take a hedging strategy is available. However, since there is no extra benefit to completing both projects, there is still a decision to be made into which project to back. Given the identical probability of successfully completing each project, the optimal strategy (Figure 22) is now to invest in both projects together, to see which succeeds.

Table 19: A more relaxed revenue function does not penalise the hedging approach and so allows combinations in which **stage 1** of the non-successful project was completed successfully.

Scenario	1	2	3	4	5	6	7	8	9
Project 1	0	1	2	0	1	2	0	1	2
Project 2	0	0	0	1	1	1	2	2	2
Revenues (£B)	0	0	20	0	0	20	20	20	20

5.4.3 *Under Perfect Additivity (Independent Projects)*

The above strategies hold when the project revenues are mutually exclusive. No extra benefit is obtained from taking both projects through to successful completion, and so a trial-and-error approach is taken to first resolve the most uncertainty, and then invest after determining the most profitable (i.e. likely) result.

If the projects rewards are independent from one another, i.e. the rewards are additive, then the pattern of investment takes on a portfolio-type effect, where investment in both projects occurs at the earliest **stage**, with successive investments occurring in outcomes in which the **research projects** are successful.

Project 1, Project 2

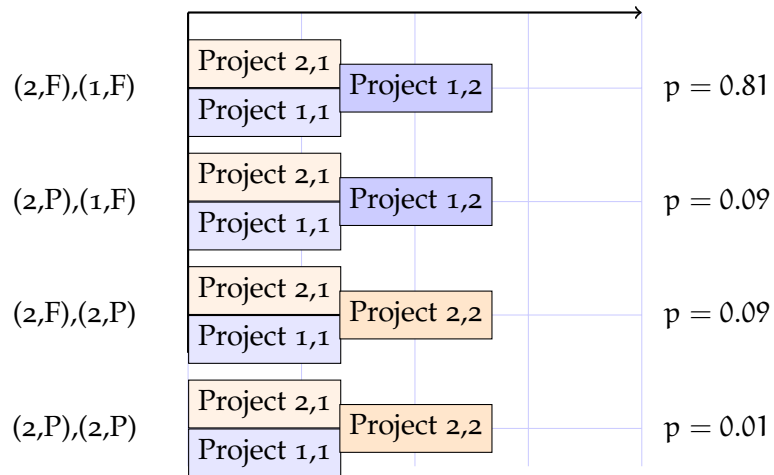


Figure 22: Optimal strategy under mutually exclusive revenue function with hedging

Table 20: Under perfect additionality, revenues are summed, so successfully completing both projects (scenario 9) results in a doubling of revenue over completing just one project successfully.

Scenario	1	2	3	4	5	6	7	8	9
Project 1	0	1	2	0	1	2	0	1	2
Project 2	0	0	0	1	1	1	2	2	2
Revenues (£B)	0	0	20	0	0	20	20	20	40

Project 1, Project 2

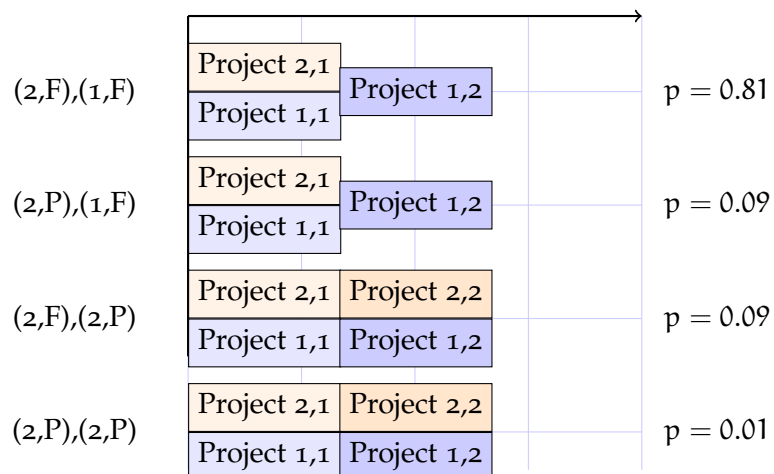


Figure 23: Optimal strategy under perfect additionality

5.4.4 Imperfect Additionality (Partially-Independent Projects)

Finally, if additionality is imperfect, or the reward for successfully completing both projects is less than the sum of the reward for completing each project individually, there are diminishing marginal returns to investment. This is the case in the results outlined in the Chapter 6. In the example shown here, the reward for successfully completing both projects together is £30B, as opposed to £40B under perfect additionality, or £20B for either project on its own.

Table 21: Example revenues under imperfect additionality

Scenario	1	2	3	4	5	6	7	8	9
Project 1	0	1	2	0	1	2	0	1	2
Project 2	0	0	0	1	1	1	2	2	2
Revenues (£B)	0	0	20	0	0	20	20	20	30

According to Shittu et al., (2010), under diminishing marginal returns as a function of investment in multiple R&D projects, i.e. where additional revenue from successful completion of each successive project is a decreasing function of the number of projects, the observed strategy is a portfolio type approach in a mixture of projects that are most independent. In the simple example presented here with just two (now independent) projects, the optimal strategy reflects this, with early investment in both projects to increase the likelihood of successfully completing either or both.

The optimal strategy depends on the ratio of the cost of conducting the second *stage* of the project versus the marginal reward gained. If the expected cost/benefit is greater than zero, then the strategy will be identical to that shown in Figure 23. Otherwise, the strategy is to only invest in the 2nd *stage* of project 1 if the first *stage* of project 2 fails, as shown in the previous Figure 22. Note that this is the same solution as in that under the less restrictive example in Section 5.4.2. For the example shown, the threshold is \approx \$25.9B.

5.4.5 Extension to Probabilities <1

The situation becomes less straightforward once we implement more realistic probabilities for each of the *stages* of the *research projects*. Now, the entire suite of nine outcomes are included in the computation of the expected net-present value of the research portfolio. Again, if the projects are mutually exclusive, then the strategy is the same as in the simplest case — the *research project* with the lower probability of successfully completing the earlier *stage* is carried out first.

In cases where the revenue is received if either project 1 or 2 succeeds, then a slightly different strategy is seen. In this case, both [stages](#) of the Project 2 are completed before embarking on Project 1, and then only in cases when the result was (1,F) or (2,F) for Project 2 (outcomes 1 through 6). This strategy works because the revenues received in outcomes 3 and 6 outweigh the extra cost of i) investing in Project 2, [stage 1](#) earlier to allow for investment in Project 1 by 2050 and ii) investing in Project 1 at all. Note also that this change in strategy is sensitive to the probability differential between [stages 1](#) and 2 of the two [research projects](#). As soon as the probabilities of success become too small, the expected revenues of outcomes 3 and 6 shrink to the point that it is not worth the extra cost of investing in Project 1.

5.5 SUMMARY

In this chapter, I presented a conceptual discussion of [research projects](#) to model the process of learning; reducing uncertainty over time. I explained how the use of [research projects](#) broken into sequential [stages](#) requires modelling of endogenous uncertainty — where the decision variables in the model alter the timing of the observation of uncertainty. While the probabilities associated with the likelihood of a [research project](#)'s success are specified exogenously and do not change as a consequence of the decision to invest in a [research project](#), the timing of the observation of the research results are decision dependent. This means that a special model formulation is required.

In this thesis, an existing mixed-integer stochastic programming model is used to allow the binary logic associated with switching on-and-off [NACs](#) in an exogenously defined [scenario tree](#) of the [outcomes](#). This model is linked to the [ESME](#) model via a revenue-function, which represents the added value of successfully completing a specific combination of [stages](#) by a specific key-year. For each combination of [stages](#) and key-years, a single [ESME](#) scenario is run, implying that there is a natural constraint to the size of the problem that can be achieved through the time it takes to run the [ESME](#) model. The formulation of the learning model allows one or more [projects](#), [stages](#) and key-years to be defined, although the results in the next chapter present the findings for one or two [research projects](#), with one or two [stages](#) and up to four key years.

Running the learning model in isolation, using a dummy revenue-function, reveals some of the behaviours expected of such a model. I showed how the model front-loads uncertainty, preferring a cheap but probable failure to a expensive but unforeseen failure, and hedges against the failure of projects by investing in multiple projects in parallel, exhibiting opportunistic investment in winners after sifting out losers.

LEARNING UNDER DYNAMIC UNCERTAINTY: RESULTS

In this chapter I present the results from the linkage of the learning model to the [ESME](#) model. In Section [6.1](#) I describe the case studies used to demonstrate the learning model and the link with the [ESME](#) model. In Section [6.2](#), I present the learning model results with examples of [research projects](#) of one phase - both for biomass availability and for CCS build rate. I then explore the combination of the two projects and the degree to which the projects complement or interact with one another (through the [revenue function](#)). The interactions are investigated through looking at one- and two-[stage research projects](#) with revenue streams that are received only at the end - a “big-pharma” style of structuring research, or with interim revenues - “partial-revenues”. The results cover the different strategies generated by the model in response to the input conditions which reveal some counter-intuitive patterns. The use of two-[stage research projects](#) increases the flexibility available to the model. The degrees of freedom within the learning model necessitate an exploratory approach, so Monte Carlo sampling is used to explore the model input space. This gives insights into the trade-offs across and between different optimal strategies under various combinations of inputs. I investigate a selection of the strategies presented by the model. In Section [6.3](#), the results of the model are presented for the investment case into the two [research projects](#) — biomass availability and CCS build rate. Finally, in Section [6.4](#), I outline the [ESME](#) results in the context of the combination of parameter inputs used in the learning model results.

In Section [5.2](#) on page [104](#), I clarified the terms [outcome](#), [stage](#) and [scenario](#). I now add to these terms the concept of a STRATEGY. A [strategy](#) is the collection of investment decisions¹ made by the learning model as a consequence of the inputs presented to the model. Essentially, the [strategy](#) encapsulates the results from the learning model in a human-readable form, and enables a higher-level interpretation. The investment decisions are a function of the outcomes generated from the combination of available [stages](#). In the results shown later, there are nine outcomes which enumerate all the possible pass/failure of the pair of two-[stage research projects](#). These outcomes are weighted by probabilities. The investment decisions are simultaneously influenced by the [ESME](#) scenarios which comprise the revenue-function, and quantify the value of the timing of the suc-

¹ In technical terms, the strategies are compiled from the decision variable X_{ijts} which represents the decision to invest in project i , [stage](#) j , in time period t and [outcome](#) s . See Appendix [B](#) for the model formulation.

successful [outcome](#) of the [research projects](#). The strategies, represented by plots showing the timing of investments under the various outcomes, illustrate patterns of investment which are strategic in nature and provide insights into sequential investment under uncertainty.

6.1 CASE STUDY

In this section I describe the case studies used to demonstrate the learning model and the link with the [ESME](#) model. In Chapter 5 I outlined the formulation of the learning model and the linkage to the [ESME](#) model. The choice of R&D projects in the learning model is linked to a [revenue function](#) representing the added value of completing a project, which is computed using the [ESME](#) model.

It is essential to identify the most important uncertain parameters to both limit the number of scenarios run in the [ESME](#) model, and because of the computational cost of large numbers of [research projects](#) in the learning model due to the multi-stage stochastic formulation. In Chapter 4 I identified these parameters.

In Section 6.1.1 I discuss some of the nuances in the linking of the two models. In Section 6.1.2, I devote individual sections to each of the two selected parameters, the availability of domestic biomass, and the build rate of CCS technologies.

6.1.1 *Choice of Parameters*

In Chapter 4, the two most influential [ESME](#) parameters amenable to a learning approach are found to be the availability of domestic biomass and the build rate of CCS technologies. An investigation of both of these parameters is interesting because neither are associated with technology costs, a focus of previous energy system studies (such as Barron et al., 2015; Bosetti et al., 2009).

The sensitivity analysis showed that there is a large and non-linear financial benefit associated with the build rate of CCS, and the results from the [ESME](#) model in Section 6.4 show that this is exaggerated if biomass is unavailable. Similarly, the availability of domestic biomass is even more influential, particularly if CCS is unavailable. There is an interesting interaction between the two parameters which is not easy to determine in isolation, and both parameters have repercussions across the energy system and over time that are different from one another.

Clearly, the resolution of these critical uncertainties is of key importance. The financial benefit of the timing of the resolution of these uncertainties is quantified using the [ESME](#) model. In addition, the costs and other aspects of performing [research projects](#) to resolve these critical uncertainties could be very different. Furthermore, given that the successful deployment of one of these critical decarbonisa-

tion options is likely to interact with the value of the other, the optimal investment [strategy](#) in resolving the combination of uncertainties would be expected to reflect the findings in Section [5.4.4](#).

6.1.2 Defining the Case Studies

In Section [5.1.1](#), I outlined the use of *research projects* as a proxy for the learning process related to resolving the uncertainty associated with a parameter. The words research and learning are used interchangeably.

Four high-level cases are used to explore the prospects for learning about the two uncertain parameters. These are outlined in Table [22](#). The cases are structured over three dimensions:

1. the number of [stages](#) in the [research projects](#)
2. whether the [research projects](#) are modelled together or independently
3. how the revenues are distributed over the [stages](#)

The one-[stage](#) cases simplify the interpretation of the model results, allowing an exploration of the [research projects](#) with a large effect, while excluding the complicating effects of partial revenues. The cases with two-[stage research projects](#) demonstrate the value of the flexibility to abandon unsuccessful projects, and are explored using both big pharma and partial revenue distribution of revenues.

Under the big-pharma style of revenue distribution, all the [stages](#) of a project must be successfully completed to receive the reward. Under the partial revenues distribution of revenues, the revenues are divided according to the results from the [ESME](#) model.

Table 22: The cases and location of the results in this chapter

stages	Complexity	Big Pharma	Partial Rev's	ESME Results
1	Individual	6.2.1	N/A	Omitted
1	Combined	6.2.2	N/A	6.4.3
2	Individual	Omitted	Omitted	6.4.1 & 6.4.2
2	Combined	6.3.1	6.3.2	Omitted

The following sections discuss the values assigned to the [ESME](#) model parameters linked to the [stages](#) of the [research projects](#) in the learning model.

6.1.2.1 Availability of Domestic Biomass

In the one-[stage](#) cases, the availability of domestic biomass increases over just two levels, shown in Table [23](#). In the two-[stage](#) cases, the

availability of domestic biomass increases over three levels as shown in Table 24.

As the model is run in 10 year time-steps, up to 1800 TWh can be consumed over one time period, if the research and development process is successful.

Table 23: Research levels for domestic biomass in the one-stage case

Level	Availability (TWh/year)
0	0
1	180

Table 24: Research levels for domestic biomass in the two-stage case

Level	Availability (TWh/year)
0	0
1	90
2	180

These quantities agree with the range presented in Department of Energy and Climate Change, (2011). The values increase linearly in the levels reflecting the possibility of bringing more land into production of domestic biomass, through the cultivation of energy crops or diversion of bio-waste streams.

6.1.2.2 Build Rate of CCS Plant

The build rate of CCS plant, both generation and industrial increases over two or three levels as shown in Table 25 and Table 26 respectively. As the model is run in 10 year time-steps, up to 20GW of CCS plant can be built in one decadal time period, if the research and development process is successful.

Table 25: Research levels in the one-stage case

Level	Build Rate (GW/year)
0	0.015
1	2.000

These rates are comparable to historical deployment rates for novel technologies in the UK. Kramer et al., (2009) discuss the physical, policy and fiscal challenges associated with transforming an energy system. Build rate constraints are used within an energy system model

Table 26: Research levels in the two-stage case

Level	Build Rate (GW/year)
0	0.015
1	1.000
2	2.000

to approximate these external constraints. So while a [research project](#) (or equivalent investment in learning) may be able to increase the build rate above zero, there is likely to be a separate upper bound on the level of deployment that is achievable. What is considered achievable is debatable and is an assumption within this framework. Sustaining the maximum value of 2GW per year for a decade would require a significant effort, and would match the peak of the UK's dash-for-gas or exceed the build out of nuclear power stations in France. The build rate in level zero reflects the construction of a pilot plant of 150MW over a ten-year period which at the time of the study was mandated in UK energy policy.

The specific value of the interim [stage](#) in both of the two-stage [research project](#) cases is an important, but not a critical assumption. While the assumption of a linear step in values does influence the resulting [revenue function](#), the learning model also takes into account the cost and probability of success associated with each [stage](#). For example, if the successful completion of the first [stage](#) of a two-stage [research project](#) resulted in a CCS Build Rate of 0.5 GW/year instead of 1.0 GW/year, then one would expect the revenue to be lower. Given the non-linearities associated with both of the parameters (from the results of the sensitivity analysis in Chapter 4) it is likely that this could be much less. However, the cost of conducting the associated [research project](#), and the probability of success of the project would also be different. A reasonable assumption could be that the first [stage](#) would cost less, or be more likely to succeed. These assumptions on costs and probabilities are tested in the results Sections 6.2 and 6.3.

On the nature of exactly what learning is going on, for this parameter, the role for blue-sky research is relatively small in the scaling-up of a nationwide programme into CCS. There are a range of things that could influence the build-rate of CCS, including availability of labour, manufacturing ability and supply-chain for plant components, planning laws and so on. And so rather than an individual [research project](#) in the traditional sense of the word, the learning process modelled encompasses all of these details into one bundled uncertainty. For an estimate of the value of the uncertainty, this aggregate treatment is deemed acceptable. This shortcomings mean that little insight is available at a higher resolution, such as discerning the value of supply chains versus labour.

6.1.3 Revenue Function: Partial Revenues

To generate the [revenue function](#) for the learning model with two [stages](#), partial benefits and two key-years, there are 36 scenarios to run in the [ESME](#) model. Throughout the results, the number of key-years is restricted to two, 2030 and 2040, as the results from the [ESME](#) scenarios in [Section 6.4](#) indicate that the early resolution of uncertainty has little effect upon the results. For several of the scenario runs, the model was unable to solve due to issues with the copying of the non-anticipativity constraints. This may have been due to a bug in the model, or an error in the formulation. The formulation constrains the decision variables only for the earlier periods so the model cannot anticipate an unresolved uncertainty. The model failed when the decision variables were constrained most heavily i.e. out to 2030, and with significant successful [research projects](#). One possible reason is that this constraining of early decisions locked in particular patterns of investment throughout the remainder of the time horizon which meant the model was unable to meet certain operational constraints.

As a work-around, I performed a multiple-regression to interpolate the 6 missing values of the [revenue function](#) from the 30 values that were available, using the combination of levels of the two projects as the independent variables. The results from the regression analysis show that the [research projects](#) are more valuable in 2040 than in 2030. However, this omits the correlation between the time periods due to project irreversibility. The intercept can be interpreted as the average extra value to the energy system of the successful completion of any [stage](#) of either of the [research projects](#) (around £7bn). Early success in either of the [CCS stages](#) is more valuable than early completion of the biomass project, while the values are similar, with [CCS](#) slightly higher for a successful 2040 completion.

[Table 37](#) in [Appendix G](#) on page [222](#) shows the interpolated values. Given the high value of R^2 , the interpolated values seem a relatively good estimate of the true values. Note that the response of the revenue from [ESME](#) is almost linear in the changing levels, although the statistical model tends to over-estimate the benefit of completing both the projects to [stage 2](#). This is an example of the non-additivity discussed in [Section 5.4.4](#). Comparing the actual and predicted values show a small discrepancy in the scenarios 2 to 15². The values for scenarios 11 and 26 were therefore increased slightly.

The values of the revenue function (£18-74bn) represent a range of between 4% to \approx 17% of the discounted total energy system cost. Note that this range should be treated with caution, as it is a function of what is included in the cost equation of the [ESME](#) model. For

² Scenario 1 is omitted from the table as it is the null case - no action is taken and the revenue is therefore 0

Formula: $Y \sim \langle \text{Bio 2030} \rangle + \langle \text{Bio 2040} \rangle + \langle \text{CCS 2030} \rangle + \langle \text{CCS 2040} \rangle + \langle \text{intercept} \rangle$

Number of Observations: 29

Number of Degrees of Freedom: 5

R-squared: 0.9891

Adj R-squared: 0.9873

Rmse: 1.7687

F-stat (4, 24): 544.3978, p-value: 0.0000

Degrees of Freedom: model 4, resid 24

Variable	Coef	Std Err	t-stat	p-value	CI 2.5%	CI 97.5%
Bio 2030	1.92	0.54	3.54	0.0017	0.86	2.98
Bio 2040	13.63	0.49	27.89	0.0000	12.67	14.59
CCS 2030	4.07	0.49	8.32	0.0000	3.11	5.02
CCS 2040	14.40	0.53	27.28	0.0000	13.36	15.43
intercept	6.86	0.95	7.24	0.0000	5.00	8.72

Figure 24: Results of the multiple regression

example, the total energy system cost includes the non-energy related costs associated with transport and buildings.

6.1.4 Revenue Function: Big Pharma

Once the [revenue function](#) for the partial-revenues had been computed, it was possible to generate a [revenue function](#) for the “big-pharma” results, in which the project was split into multiple [stages](#), but there was no interim revenue generated from the first [stage](#). In fact, the project would need to be taken through to successful completion of the final [stage](#) to earn the revenue. The [revenue function](#) was edited to set the value for those scenarios in which only [stage](#) one was completed to the revenue of the equivalent [ESME](#) scenario. For example, the revenues for scenarios 2 and 3 were set to zero (no projects completed to [stage](#) 2). Scenario 5 is set to the same value as that for 6 - no extra revenue is earned for having [stage](#) 1 of the bio project completed in 2030. The final [revenue function](#) is shown in [Table 38](#) in [Appendix G](#) on page [223](#). The big-pharma revenue function covers a range of between £36bn and 72bn (8% to \approx 17% of the total energy system cost).

Table 27: The probabilities of the research project outcomes in the learning model. The probabilities for each column sum to one (values are rounded to two decimal places). These probabilities are ‘invented’ to demonstrate the learning model and bear no relation to the probability distributions contained in the [ESME](#) data set.

Outcome (Bio), (CCS)	Probability of stages (stage 1 , stage 2)			
	(.25, .80)	(.40, .50)	(.50, .40)	(.80, .25)
1: (1,F), (1,F)	0.56	0.36	0.25	0.04
2: (2,F), (1,F)	0.04	0.12	0.15	0.12
3: (2,P), (1,F)	0.15	0.12	0.10	0.04
4: (1,F), (2,F)	0.04	0.12	0.15	0.12
5: (2,F), (2,F)	0.01	0.04	0.09	0.36
6: (2,P), (2,F)	0.01	0.04	0.06	0.12
7: (1,F), (2,P)	0.15	0.12	0.10	0.04
8: (2,F), (2,P)	0.01	0.04	0.06	0.12
9: (2,P), (2,P)	0.04	0.04	0.04	0.04

6.1.5 Learning Model Outcomes

Table 27 lists the probabilities of the nine [outcomes](#) over the four combinations of first- and second-stage probabilities presented in the results that follow.

Note that while the probability of [outcome 9](#) (both projects successfully completing to [stage 2](#)) stays constant across the four combinations of probabilities, the probabilities of the other outcomes change as the balance of probabilities move from front-loaded, to back-loaded. Also, the changes in the probabilities are not symmetrical, due to the dependency structure (i.e. the consecutive nature of the projects) of the [scenario tree](#). For example, when front-loading the uncertainty (the first column), the most likely result is [outcome one](#) - the failure of both the projects - at 0.56. In contrast, the back-loading of uncertainty (the rightmost column), the probabilities of the outcomes exhibit a more uniform distribution, with the most likely result the failure of both projects after the second [stage](#) - at 0.36.

These probabilities are used to weight the discounted costs and revenues in the learning model and are important to understand why particular investment strategies evolve from the model inputs.

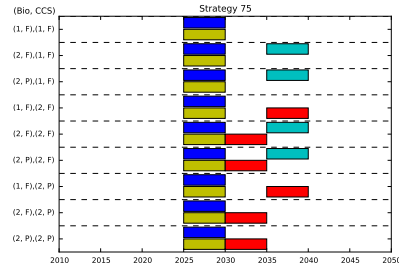
6.1.6 Investigating Strategies

In this section I introduce and discuss some of the more complex strategies exhibited by the learning model in the results that follow.

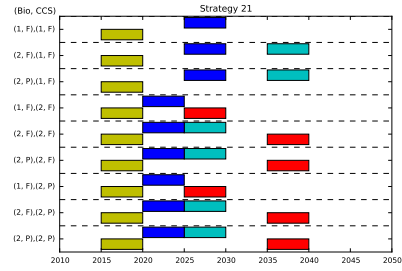
No insights into the likelihood of the occurrence of these strategies, or their being the optimal choice can be made, as they were generated by using a Monte Carlo sample to stimulate the research project costs to explore the output space of the model. However, it is possible to discuss under what conditions these strategies occur. The model generates a very large number of distinct strategies, and an exhaustive discussion of every [strategy](#) would offer diminishing insights. The number of strategies is a function of the degrees of freedom available to the model, such as the number of projects, [stages](#), key years and revenue structure. For starters, there is little gained from discussing the inverse of the strategies (merely swapping one [research project](#) for another), and the most basic strategies, where investments are only made into one [stage](#) of a research project regardless of the results, are self-evident. The following therefore concentrates on some of the more complex strategies, despite the fact that they could be considered special or corner cases.

Figure 25 shows the selected example strategies. Each of the charts displays the time horizon of the learning model in 5-year time-steps. On the y-axis, the nine possible outcomes are shown. The key to these outcomes are shown as labels to the y-axis, with the result of the biomass project on the left, and the results of the CCS project on the right. For example, (2, P), (1, F) means that in this [outcome](#), the biomass project succeeds ([P]asses) through to the final (second) [stage](#), while the CCS project fails after the first [stage](#). The coloured blocks represent the decision to invest in a particular [stage](#) of a [research project](#). The blue and cyan blocks represent the first and second [stages](#) of the biomass availability [research project](#). The yellow and red blocks represent the first and second [stages](#) of the CCS build rate [research project](#).

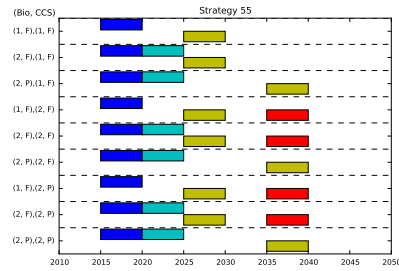
Strategy 75, shown in Figure 25a shows a hedging [strategy](#) in which there is investment in the first [stage](#) of both the biomass and CCS projects in 2025. This means that the projects are completed in 2030 (all projects are assumed to last five years), and the uncertainty surrounding which projects will succeed is therefore resolved in 2030. In the first three outcomes, the CCS project fails, and so the recourse is to invest in the second [stage](#) of the biomass project. In the fourth and seventh outcomes, the biomass project fails after the first [stage](#), and the recourse is to invest in the second [stage](#) of the CCS project. In the fifth and sixth outcomes, neither project fails after the first phase, but the CCS project fails after the second [stage](#). However, this is not known until the end of the second [stage](#) of the CCS project, which commences (as in outcomes eight and nine). Upon discovering the failure of the CCS project, investment switches back to the biomass project. In the eighth and ninth outcomes, the CCS project succeeds after the second phase and no further investment takes place. This complex investment behaviour, in which different actions occur de-



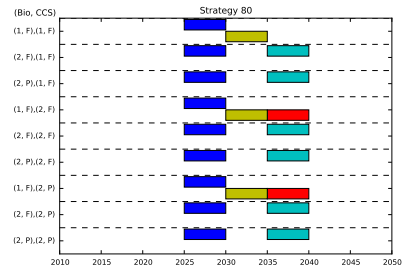
(a) Strategy 75



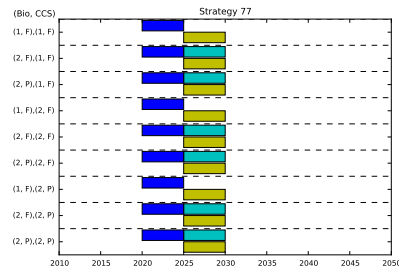
(b) Strategy 21



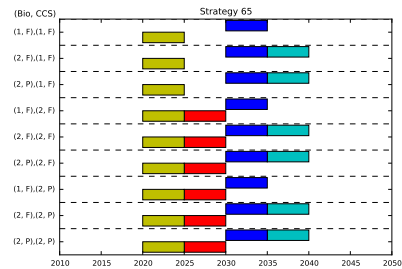
(c) Strategy 55



(d) Strategy 80



(e) Strategy 77



(f) Strategy 65

Figure 25: Some example strategies. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.

pending upon the resolution of [stages](#) of uncertainty is typical of the results from the learning model. In this [strategy](#), early investment in the first [stages](#) of each of the projects occurs so that rapid investment can take place in whichever project is successful.

In Strategy 55, shown in Figure 25c, we see very early investment in the biomass availability project, despite there being no return until 2030. This is to allow a chain of investment, continuing with the second [stage](#) of the biomass project, and under outcomes where the biomass project fails, bringing forward investment in the first [stage](#) of the CCS project to 2025. This then presents the option of investing in the second [stage](#) of the CCS project which is timed to complete by 2040. In the outcomes where the biomass project succeeds, then investment in the first [stage](#) of the CCS project is taken in 2035 and completed by 2040.

In [strategy](#) 21, shown in Figure 25b we again see early investment in 2015, but this time in the CCS project. If the project fails at the first [stage](#), investment in the biomass project is delayed until 2025, so that the revenue is recouped in 2030 upon the success of the first [stage](#) of the project. If the biomass project succeeds in the first [stage](#), investment in the second [stage](#) continues in 2035. If the first [stage](#) of the CCS project succeeds, in this [strategy](#), the investment switches tack, investing in the first phase of the biomass project in 2020. If this investment succeeds, investment continues into the second [stage](#), completing in 2030. If the first [stage](#) of the biomass fails, then investment switches back to the second [stage](#) of the CCS project. In the situation where both projects are successful at the first [stage](#), as in outcomes five, six, eight and nine, the investment in the second phase of the CCS project is delayed to 2035 and completed in 2040.

Strategy 77, shown in Figure 25e demonstrates unconditional investment in the first-[stage](#) of the CCS project in 2025 (collecting revenue in 2030), while investment in the biomass project takes place in 2020, continuing onto the second [stage](#) in 2025, also with the intention to collect the revenue in 2030. This is an example of a situation where the second [stage](#) of the CCS project is too expensive to consider, while the first [stage](#) is a no-regret investment. This is also a case where an identical [strategy](#), shifted ten years later, is also plausible.

The resolution of the [revenue function](#), computed from the ESME model has a direct effect upon the strategies. The use of two key years, 2030 and 2040, means that the investment in the [research projects](#) revolves around these two thresholds. A positive aspect of this is that it makes interpretation of the results easier. If there were many key years, it would become increasingly difficult to discern whether the particular timing of an investment occurs due to an earlier revenue, or due to the need to make early investment decisions to obtain a larger reward in a later period.

6.2 ONE-STAGE RESULTS

In this section I present the results from the learning model. The structure follows an investigation into biomass availability and CCS build rates individually, followed by the combination of the two.

As the strategies given by the learning model are potentially numerous, the presentation of results are initially kept deliberately simple, gradually building in complexity. In the later sections, the degrees of freedom available to the learning model generated hundreds of distinct investment strategies.

The individual one-stage research projects shown in Section 6.2.1 are much simpler than the combination of projects, and give just three potential strategies — don't invest, invest in 2020, or invest in 2030.

For the combination of one-stage research projects in Section 6.2.2, the optimal strategy takes into account the four possible outcomes (the success and failure of each project), and the trade off between investing now, or waiting to invest given the expected revenue of the research project. This gives ten viable strategies, four of which include: invest in both projects; neither of the projects; CCS first, then biomass; or biomass, then CCS.

6.2.1 Individual Projects - One-stage

Using the learning model, we can explore the optimal investment strategies under different input conditions. With just one research project of one stage, it is possible to view the outcomes for all probability/cost combinations through normalising the costs. For example, the discounted cost of a project with a 10% likelihood of completing successfully is equivalent to a project with half the value, but a 20% likelihood of completing successfully. By normalising the costs of projects and the total cost of the research pipeline, where the cost of the project and objective function are divided by the probability of the projects' success, we can present the strategies for all project cost/likelihood combinations using one line, as shown in Figure 26.

Plotting normalised cost of the research project against the normalised objective function of the learning model³ the resulting line is a convex piecewise function. Each segment of the piecewise objective function represents a different stochastic investment strategy. The vertices where the lines meet, are the points at which the model switches to a different strategy. The normalised cost at which the vertices are positioned are the values which determine the choice of the strategy.

The gradient of the lines represent the rate of change of the value of the research pipeline as the project costs increase. The origin of the

³ the discounted value of the research pipeline divided by the probability of the research project succeeding

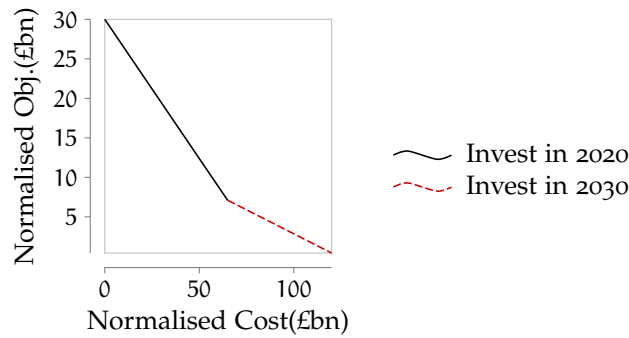


Figure 26: The normalised objective function of the learning model given the change in normalised cost of the biomass availability project with a 10% discount rate, and with an artificially extreme decrease in revenue - £30bn in 2030 and £15bn in 2040

line on the y-axis is £30bn, the value of the research project finishing in 2030 when projects costs are zero.

The position of the vertex which joins the two strategies (invest in 2020, or invest in 2030), is a function of the balance between the revenues obtained from successfully completing the project in 2030 and 2040, and the discount rate. The discount rate only affects the project costs, as the revenues are already discounted to 2010 by the ESME model. Hence, a project's normalised cost increases until the point where it is more cost effective to delay the project to a later [stage](#), despite the change in revenue (assuming that less revenue is gained if biomass availability is delayed).

In the exaggerated example shown in Figure 26, for any normalised cost less than £60bn, the optimal [strategy](#) is to invest in a bio availability [research project](#) in 2020, with the result determined in 2030 ([research projects](#) are assumed to last 10 years). When expected cost is greater than £60bn but less than £120bn, the optimal [strategy](#) is to invest 10 years later. As the normalised cost gets closer to the maximum of £120bn, the expected net present value of conducting the research decreases to zero. Above £120bn, the optimal [strategy](#) is not to invest.

In the actual case, the loss of revenue from delaying biomass availability to 2040 is just 10% of the 2030 value - around £3bn. Holding CCS build rate to zero, the value of biomass availability is £30bn in 2030 and £27bn in 2040 (discounted to 2010). Using a social discount rate of 3.5%, the convexity becomes less pronounced, as the change in cost from delaying has less effect upon the drop in revenue. This means that the point at which the model switches between strategies is in negative cost - the only optimal [strategy](#) is to delay investment to 2030, or to not invest (depending on the cost of the project).

Holding domestic biomass availability to 2.2 TWh, the value of CCS is quantified by the [ESME](#) model as £34bn in 2030 and £28bn in 2040 (discounted to 2010). The behaviour of the learning model is virtu-

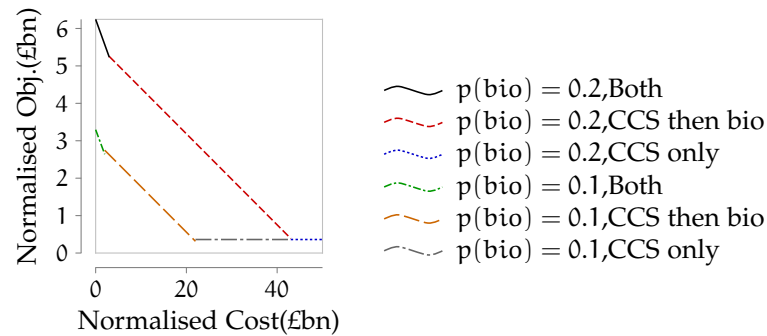


Figure 27: The normalised objective function of the learning model given the change in (undiscounted) cost of the biomass availability project with a 10% discount rate, while a CCS build rate project is fixed at 10% likelihood and a high cost of £20bn

ally the same as for the one-stage biomass project. There are just three strategies - invest in 2020, 2030, or do not invest.

6.2.2 Both Research Projects - One-stage

Combining the two research projects, the results from the ESME model give values of between £27bn and £59bn.

If we again plot normalised cost of one project against the normalised objective function, while holding the other project at fixed values, we obtain a convex piece-wise function. Unlike the function in Figure 26, this function has three segments, as the interactions between the research projects result in a more complex transition between optimal strategies. This new function is shown in Figure 27. When the normalised cost is less than £2bn, the optimal strategy is to invest in both projects in parallel in 2030. Between £2bn and £22bn, the optimal strategy is to first invest in the CCS project, in 2030, and then only invest in the biomass project if the CCS project fails. If the biomass project costs more than £22bn, then the optimal strategy is to only invest in the CCS project. Note that in this example, the cost of the CCS project is relatively high in comparison to the range of revenues and so exhibits a crowding-out effect. When the CCS project is cheaper, then the range of costs over which it is optimal to invest in both projects increases in size.

A further complication arises as now changing the probability of success of the biomass project results in a differently shaped objective function. For example, in Figure 27, doubling the probability of success of the biomass project from 10% to 20% results in increasing the range of costs that it is optimal to invest in both projects (from between £2bn and £22bn to £3bn and £43bn). The horizontal segments evident in both functions represent the objective function when only the CCS project is invested in. Given the maximisation ob-

jective, in the absence of alternative investment options which are ‘in the money’, the CCS project alone sets the minimum boundary that the objective can reach.

To further explore the behaviour of the dynamic learning model, I used Monte Carlo sampling with uniform distributions to permute the cost inputs for the model, over four combinations for the probabilities of the [research projects](#). Figure 28 shows the strategies that are invoked by the model under combinations of project costs and probabilities of success. By [strategy](#), I mean the combination of decision variables in the learning model across the various outcomes. For biomass and CCS, there are 16 possible strategies (computed from the product of four scenarios, two projects, two key years and one [stage](#)) + 1 null [strategy](#) (do nothing), although only ten of the strategies that are possible are viable, given the [revenue function](#) and discount rate. The viable strategies are listed in Table 28.

Within the constraints of the two one-[stage](#) projects that are modelled in this example, most of the strategies are simple. For example, in strategies 0, 2, 3 and 6 only one project undertaken. These strategies occur when the other project just isn’t financially viable (expected net present value is less than zero). In contrast, in strategies 4,5,7 and 9 both projects are undertaken. In other words, the model invests in both projects to hedge against the likelihood of one or the other failing. The two strategies, 1 and 8 that remain, are the most interesting in that they exhibit adaptive behaviour, where the investment [strategy](#) changes depending upon the result of an initial [stage](#). This reflects the results from the previous chapter in Section 5.4.1 (page 117), where optimal investment strategies adapted to the out-turn either continuing with a successful project or commencing investment in an alternative.

The plots are arranged so that the total cost of the [research projects](#) are displayed on the x- and y-axes, the strategies are identifiable in coloured areas that represent the convex hull of the pseudo-random inputs. Separate plots have been produced for combinations of probabilities of the projects at 80% or 20%.

The structure of the areas which represent the strategies shows how complex the choice of a [strategy](#) is across the ranges of costs that could be considered for the two [research projects](#). For example, in Figure 28d, the likelihood of both [research projects](#) is high. If the CCS project is less than £20bn and biomass project less than £10bn, then investment can proceed in both projects in 2020 (strategy 9). This is because despite the diminishing returns from successfully completing both projects, the low cost relative to the return means that it is viable to commence both. With biomass remaining below £10bn, and CCS increasing above £20bn but less than £40bn, the optimal [strategy](#) is to delay investing in CCS to 2030 (strategy 4). However, as the cost of the CCS project increases above £20bn, the cost of the biomass pro-

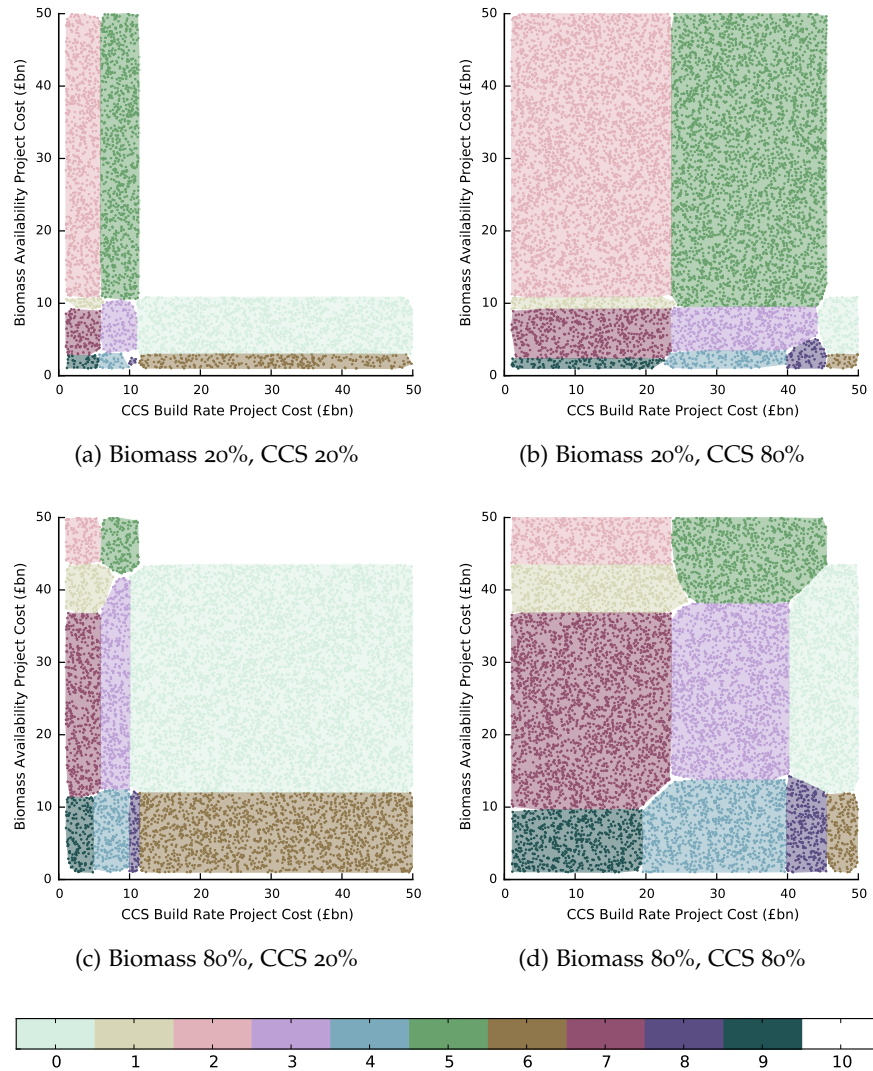


Figure 28: Four plots corresponding to probabilities of 20% or 80% for success of the research projects for biomass availability and CCS build rate

Table 28: The ten viable strategies for CCS Build Rate and Biomass Availability. The numbered key corresponds with that of Figure 28.

Key	Description
0	Biomass only in 2030
1	CCS in 2020, Biomass in 2030 only if CCS fails
2	CCS only in 2020
3	Biomass and CCS in 2030
4	Biomass in 2020, CCS in 2030
5	CCS only in 2030
6	Biomass only in 2020
7	CCS in 2020, Biomass in 2030
8	Biomass in 2020, CCS only if biomass fails
9	Biomass and CCS in 2020
10	No investment (null strategy)

ject can also increase without shifting from strategy 4. However, upon exceeding the cost threshold of £10bn (while CCS is less than £20bn) the optimal [strategy](#) switches to 7 - biomass is delayed to 2030. Increasing to £14bn (while CCS is greater than £25bn), strategy 3 is optimal - both projects are delayed to 2030.

Travelling along the x-axis, we can observe that the transition between strategies follows a pattern. Fixing biomass cost to £5bn, the transition covers strategies 9, 4, 8 and 6. In short - invest in both early, invest in both with CCS delayed to 2030, invest in biomass first and only CCS if the biomass project fails, only invest in biomass in 2020. Travelling along the y-axis, fixing CCS costs to £10bn, we observe a similar transition - moving through strategies 9, 7, 1 and 2. Again - invest in both, invest in both with biomass delayed, invest in CCS early and only biomass in 2030 if CCS fails, invest in biomass early.

Now looking at the outer edges of the plot, we can see that the strategies of last resort occur above a threshold cost. In these cases, the strategies switch to ignoring the expensive project, and focus on the cheaper alternative. In a clockwise direction these are 2, 5, 0 and 6 (invest only in CCS in 2020/2030; invest only in biomass in 2030/2020). In the top right-hand corner of each plot is the null [strategy](#) where no investments in the projects take place. The centre of the plot is dominated by strategy 3 - invest in both projects in 2030.

Looking now at Figure 28a, where the probability of success of each project is 20%, we can see that the null strategy dramatically increases in size. At the same time, the position of the strategies relative to one another remains the same, although there is some movement of the boundaries between the strategies. The conditions under which the

adaptive strategies (1 and 8) are optimal are confined to very particular combinations of biomass and CCS project costs. The extension of the backstop strategies (0, 5, 6, and 2) clearly extend out along the relevant axes, framing the area under which the null strategy is optimal.

6.3 TWO-STAGE RESULTS

In this section, I present the results of cases which are structurally the same to those presented in Section 6.2; the *research projects* have two-*stages* instead of one. This means that for each project, there is an option to abandon the project after the first *stage* if the first-*stage* fails, or, a hedging approach can be taken in which investment takes place in both projects, continuing with whichever succeeds after the first *stage*. This relaxes a major assumption of the results presented in the previous section, where the investment cost is sunk after deciding to pursue a particular project. Whereas in the previous section, the investment strategies normally switched between projects, upon the resolution of an uncertainty (normally - the project that was invested in failed), here, the decision maker may invest in the first-*stage* of project with the option of continuing or abandoning the project. The payout from the successful completion of just the first *stage* is computed according to the *revenue function* outlined in Section 6.1.3. These results give insights into projects which allow partial-revenues.

The results are split into two subsections. In Section 6.3.1, the *revenue function* is modified, as shown in Section 6.1.4 so that only the successful completion of the project to the second *stage* is rewarded. In the following Section 6.3.2, the full-*revenue function* including revenues awarded for partial successful completion of the *research projects* is used.

6.3.1 Two stage, “big-pharma”

Revisiting Figure 28a, I focus these next set of more complicated results on the lower left-hand quadrant of the plot, bounded by the total project costs of £20 bn. In this quadrant, six different strategies are present showing how the combinations of costs for the two single-*stage* projects resulted in slightly different optimal patterns of investment. Dividing both the *research projects* into two *stages*, I now vary the costs of the first *stage* only while fixing the cost of the second *stage* to £10bn, in Figure 29. Over the four subplots, we fix the probability of successfully completing each of the projects to 20%. I then show across the four plots, how changing the balance of the probabilities from the first to second *stages* changes the shape of the strategies. In Figure 30, six of the strategies common to all four plots are displayed.

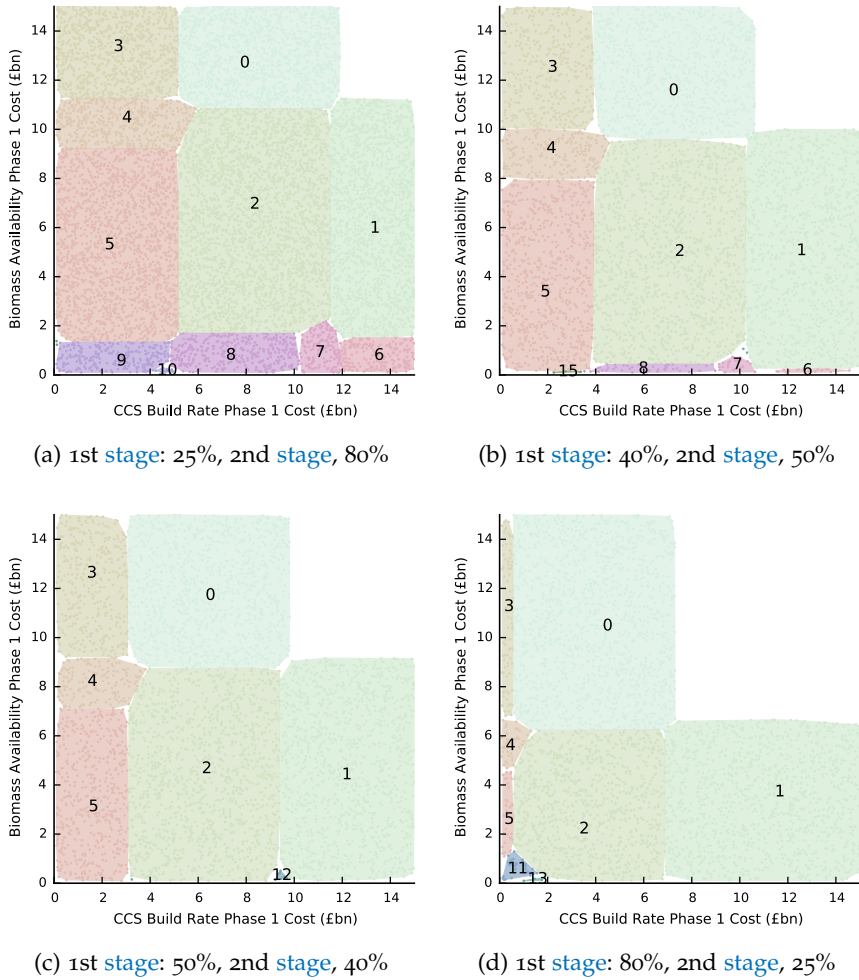


Figure 29: In these plots, the cost of the second **stage** (not included in the value on the axes) for both projects was fixed to £10bn. The cost of the first **stages** for both projects was varied up to £15bn and is displayed on the x- and y-axes for CCS and biomass availability respectively. While the probability of the project was fixed at a 20% likelihood of succeeding, the balance of the probabilities over the first and second **stages** are altered across the four plots to demonstrate the effect upon the strategies selected. The strategies numbered here bear no relation to those numbered in Figure 28.

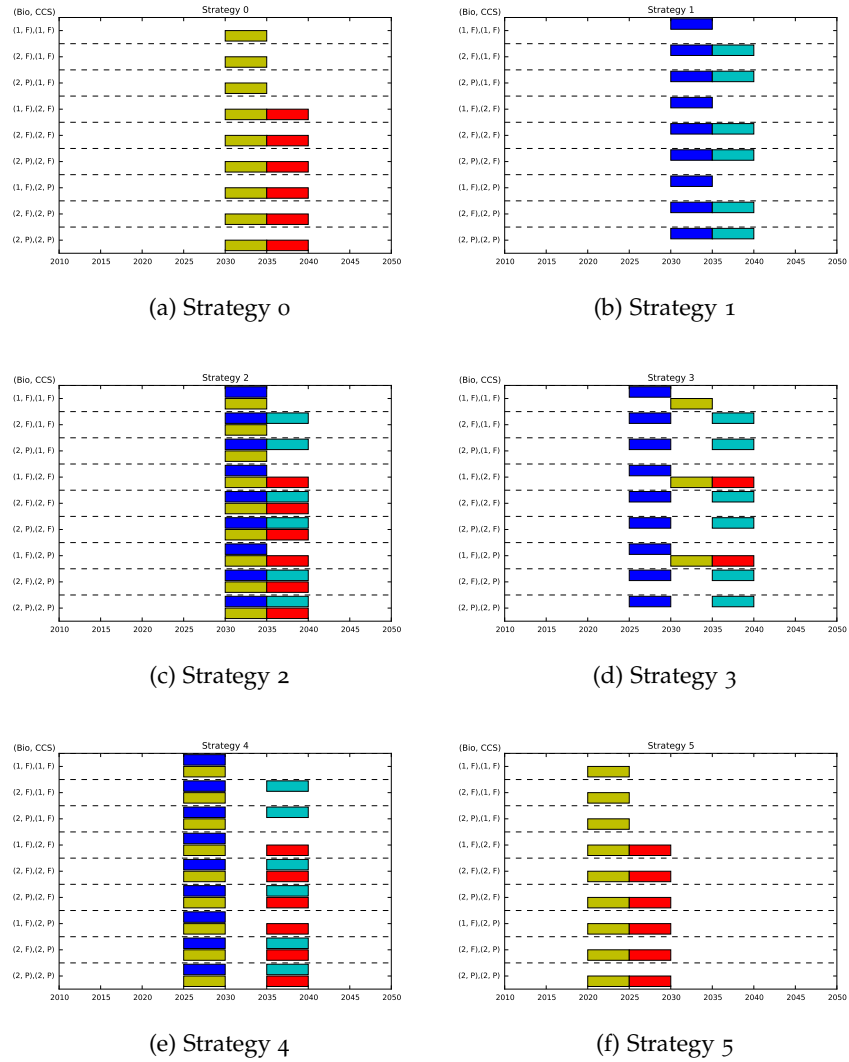


Figure 30: Six of the twelve main strategies from the example two-stage big-pharma projects, continued in Figure 31. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.

In Figure 30, fifteen strategies are incorporated within the explored decision space, an increase over that in Figure 28 due to the more degrees of freedom available to the learning model. The flexibility allowed by dividing the projects into two phases means that there are options to abandon projects after the first stage without incurring the cost of the second phase. These are listed in Table 29.

Comparing Figures 29a and 28a, the flexibility of the two-stage projects increases the range of project costs over which it is viable to invest in the projects. This is despite the probabilities of the projects' success remaining at 20%. In fact, the projects are viable over costs up to £10bn higher than for the single-stage projects. As the balance of probabilities shifts from front-loading to back-loading the uncertainty, from Figures 29a to 29d, the range of costs over which a non-null strategy is optimal declines to around £17bn. The benefit is slightly less for the biomass availability project, mainly because the successful completion of the project is less valuable alone than the CCS project and because all other costs and probabilities are held equal between the two projects, we see the result in the pattern of the strategies in these plots.

Another difference between Figures 29a and 28a is that there is a decrease in the number of strategies that are optimal within the same range of project costs. In Figure 28a there are ten strategies for project costs up to £20bn. Whereas in 29a, there are just seven projects that are optimal under £10bn (the scale on the x and y axis excludes the cost of the second stage which is fixed to £10bn). This is due to the increase in flexibility available to the learning model from being able to abandon projects which means that fewer strategies apply to greater ranges of cost combinations.

Six strategies appear in each of these plots, and are labelled from zero to five. The strategies are reproduced in Figure 30. Strategies 0, 1 and 5 focus on just one project, continuing sequentially with the project stages as the first completes successfully. Strategies 2, 3 and 4 are mixed, with investments in both projects taking place. Travelling along the x-axis as the first stage of the CCS project becomes cheaper, the strategies crossed are 1, 2 and 5. While the CCS project is expensive, investment proceeds only in the biomass project. In the middle range, investment proceeds into both projects. When the CCS project is cheap, investment only takes place in the CCS project. This latter result is perhaps counter-intuitive, because if the CCS project is cheap, there could be extra finances available to spend on a more expensive biomass project. However the marginal benefit of investing in the biomass project does not outweigh the cost and it is more cost effective to invest only in the CCS project.

In Figure 29a and to a lesser extent, Figure 29b, when the first stage of the biomass project is cheap a separate set of strategies come into play; 6, 7, 8, 9, 10 and 15. Strategy 8 is actually the direct opposite

Table 29: The 15 viable strategies for CCS Build Rate and Biomass Availability with two-stage big-pharma research project. The numbered key corresponds with that of Figure 29.

Key	Description
0	CCS ₁ in 2030, CCS ₂ in 2035
1	BIO ₁ in 2030, BIO ₂ in 2035
2	BIO ₁ , CCS ₁ in 2030, BIO ₂ , CCS ₂ in 2035
3	BIO ₁ in 2025, CCS ₁ in 2030, CCS ₂ in 2035 if BIO ₁ fails, otherwise, BIO ₂ in 2035
4	BIO ₁ , CCS ₁ in 2025, BIO ₂ , CCS ₂ in 2035
5	CCS ₁ in 2020, CC ₂ in 2025
6	CCS ₁ in 2020, CC ₂ in 2025, BIO ₁ in 2030, BIO ₂ in 2035 if CCS fails
7	CCS ₁ in 2020, CC ₂ in 2025, BIO ₁ in 2030, BIO ₂ in 2035
8	BIO ₁ in 2020, BIO ₂ in 2025
9	BIO ₁ in 2020, CCS ₁ in 2025 or CC ₁ in 2030 if BIO ₁ fails CCS ₂ in 2035, BIO ₂ in 2035
10	BIO ₁ in 2020, BIO ₂ in 2025, CCS ₁ in 2030 if BIO ₂ fails, CCS ₂ in 2035
11	BIO ₁ in 2020, BIO ₂ in 2025, CCS ₁ in 2030, CCS ₂ in 2035
12	BIO ₁ , CCS ₁ in 2020, BIO ₂ in 2025 if CCS ₁ fails CCS ₂ otherwise, BIO ₂ in 2035
13	BIO ₁ , CCS ₁ in 2020, BIO ₂ in 2025 if CCS ₁ fails, CCS ₂ otherwise, BIO ₂ in 2035
14	CCS ₁ in 2015, BIO ₁ in 2020 if CCS ₁ fails, BIO ₂ in 2025 CCS ₂ otherwise, BIO ₁ in 2030, BIO ₂ in 2035
15	BIO ₁ in 2015, CCS ₁ in 2020, CCS in 2025 if BIO ₁ fails BIO ₂ in 2025, CCS ₁ in 2030, CCS ₂ in 2035 otherwise

of 5, there is only investment in the biomass project, the biomass project is cheap and the CCS project is middling in cost. Strategies 6 (and 10) exhibit identical behaviour - investment in 2020 in CCS (biomass) followed by investment in 2030 in the biomass (CCS) with investment only occurring in the second project if the first failed. The most complex [strategy](#) occurs when the first phases of the project are at their cheapest, (less than £2bn for biomass and less than £4bn for CCS) in strategy 9. In this case, investment in the first phase of the biomass project takes place in 2020. If the first phase succeeds, investment then proceeds to the first phase of the CCS project. Investment in both second phases occurs in parallel in 2035, where possible. If the first phase fails, investment in the first phase of the CCS project is delayed to 2030, with the second phase of the CCS project following in 2035.

Travelling along the y-axis of Figure [29a](#), the results are less intuitive. While the cost of the first [stage](#) of the CCS project is less than £5bn, and as the first [stage](#) of the biomass project declines in cost, we pass through strategies 3 and 4, finishing in strategy 5. In strategies 3 and 4, investment takes place in both projects, but in 3, biomass is favoured. In the former, investment only proceeds in CCS if the first [stage](#) of the biomass project fails. In the latter, investment occurs in the first [stages](#) of both projects in parallel, with investment continuing to the second [stages](#) if at all possible. These results may be thought as counterintuitive because in this situation, the CCS project is cheaper than in all other strategies, where the CCS project is in fact favoured (compare strategies 0 and 3). Yet when the cost of the first phase of the biomass project declines, investment switches back to both projects. Traversing the x-axis, investment proceeds in both projects while biomass is less than £2bn, and the CCS project is over £10bn. However, while the CCS project lies between £5 and 10bn, investment switches entirely to the biomass project. Below £5bn, investment takes place in both projects again.

Comparing strategies 1 and 6 begins to explain this behaviour. In strategy 1, the CCS project is expensive, and the biomass project is the obvious choice given the similarity in their returns. However, below £1.5bn for the first [stage](#) of the biomass project, it becomes economical to invest in the CCS project as well - and strategy 6 comes into play. These results show that an increase in the cost of the first [stage](#) of the CCS project can result in investment in that project (comparing strategies 8 and 7 in Figures [31c](#) and [31b](#)). This shows how the interactions between these parameters are already complex even given the limited degrees of freedom available to the model.

In Section [5.4](#) of the previous chapter, the structure of the [research projects](#), i.e. the distribution of uncertainty across the [stages](#) of a multi-[stage research project](#), was shown to affect the viability of a project. In those examples, in which there were also no mid-project

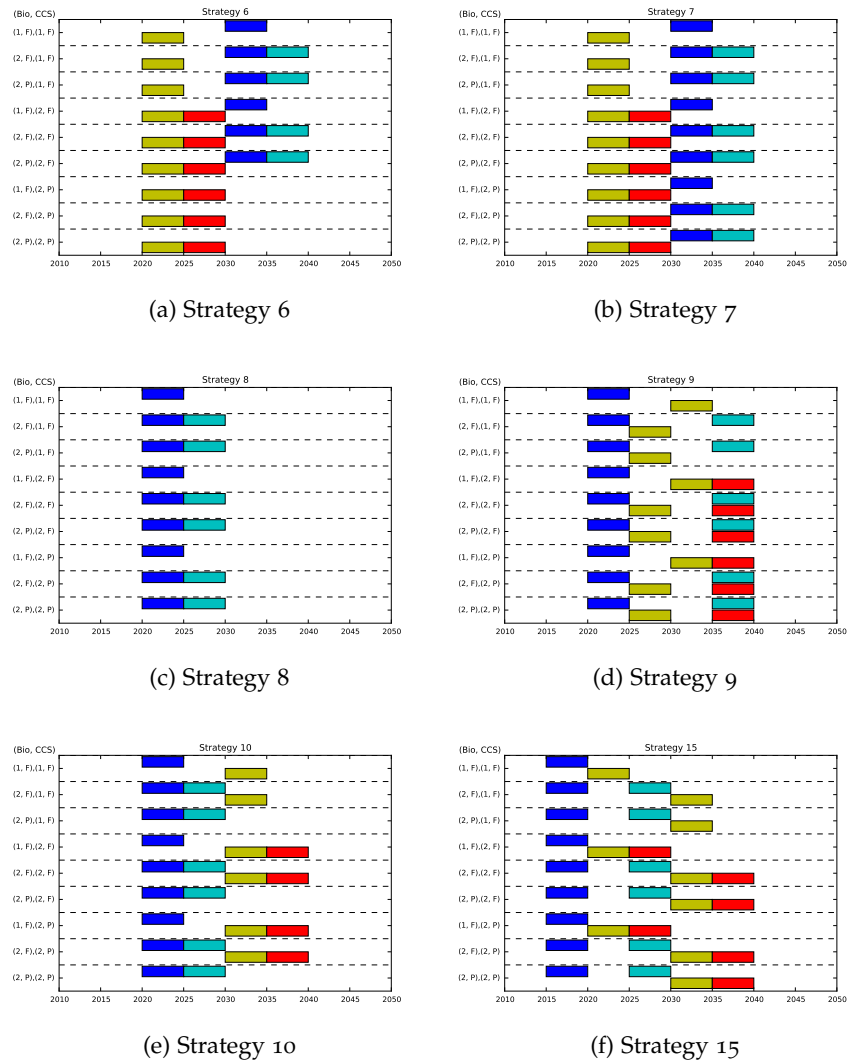


Figure 31: Six of the twelve main strategies from the example two-stage “big-pharma” projects, continued from Figure 30. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.

revenues, the projects with early resolution of uncertainty were preferred over those in which the uncertainty was resolved in a later [stage](#). This finding is also reflected in the difference between [Figures 29a and 29d](#). The cost range over which the projects are viable shrinks, while the viable strategies shrink to small cost ranges, with investments in just one project dominating the cost ranges. It seems that under the assumptions made in these plots, that there is diminished value in separating a project into [stages](#) unless the uncertainty is front-loaded.

6.3.2 *Two stage, partial-benefits*

In this section, I present the results of the cases in which the two phases of the biomass availability and CCS build rate [research projects](#) have intermediate benefits. [Tables 24 and 26](#) give the values for the input parameters, while the values for the [revenue function](#) are shown in [Table 37](#) in [Appendix G](#). As with the big-pharma results in the previous section, the combination of the two two-stage projects offers many degrees of freedom to the model. Now, with revenues after both first and second [stages](#), it is possible to extract a revenue given the extra flexibility available to the model under almost all combinations of inputs to the learning model.

The [ESME](#) parameter values associated with the two-stages of the research projects were outlined in [Section 6.1.2](#), and the resulting [revenue function](#) in [Section 6.1.3](#). While the parameter values increase linearly as a function of the [research project stages](#), the [revenue function](#) did not, and combinations of the projects further eroded linearity.

The results presented in [Figure 32](#) are displayed in the same formation as those in the previous [Section 6.3.1](#). The four plots correspond to the shift in probabilities of success of the [stages](#) from the second to first [stage](#). Six of the eighteen strategies are presented in [Figure 33](#).

The addition of partial benefits changes the pattern of strategies. Investment in only the first [stage](#) of a project is now a viable [strategy](#), and we can observe this behaviour in [Figures 33b, 33d and 33f](#). [Strategy 7](#) is the most common across the four probability combinations in [Figure 32](#). The multiple options for extracting revenue from investments results in a larger number of potential strategies over the same range of costs (for the first [stage](#)) in the case where the balance of uncertainty is front-loaded. Due to the shape of the [revenue function](#) — the first [stage](#) of the [research projects](#) result in a significant proportion of the maximum revenue obtained from completing the second [stage](#) — the shift of probability of success to the first [stage](#) no longer exhibits the penalty seen in the “big pharma” style of structuring projects. As a result, strategies which can lock in the first [stage](#) benefits, as in [Strategies 14, 15 and 16](#), are now optimal over a very wide range of costs. In fact, the behaviour exhibited matches more closely

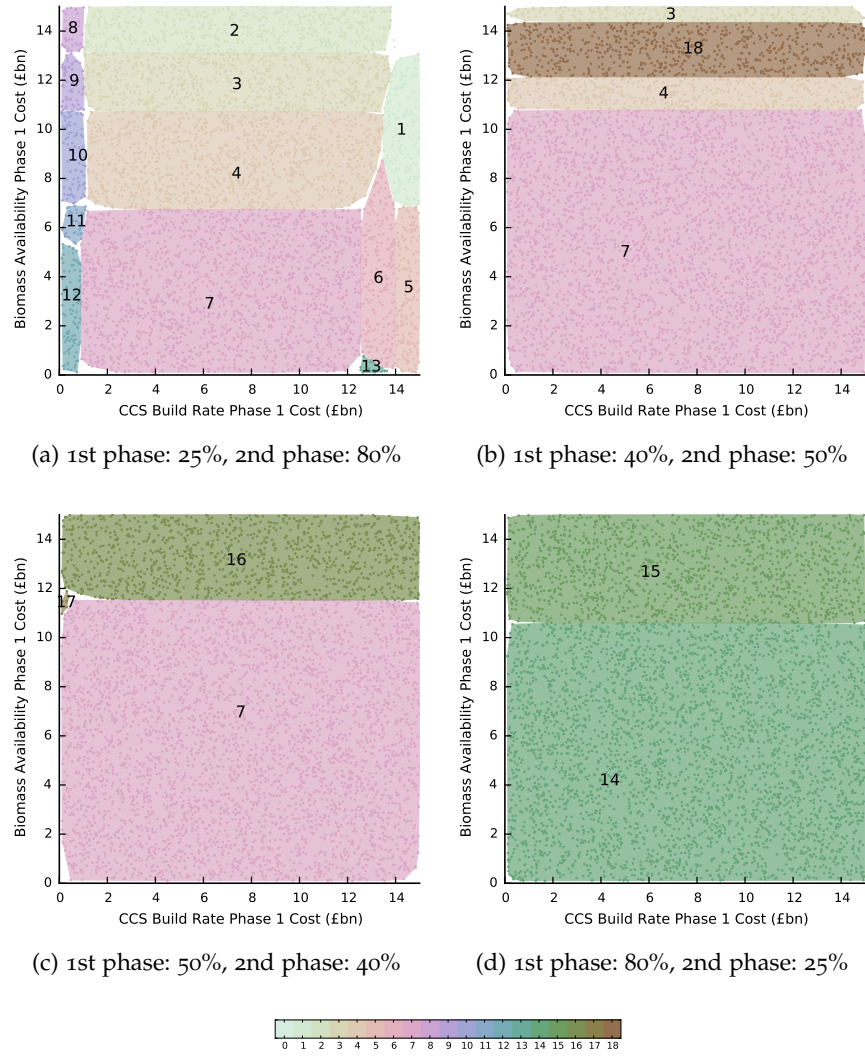
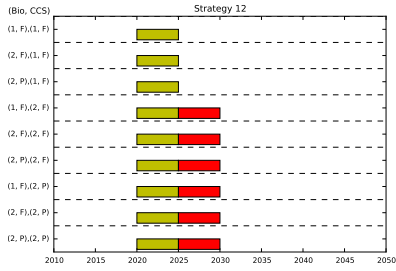
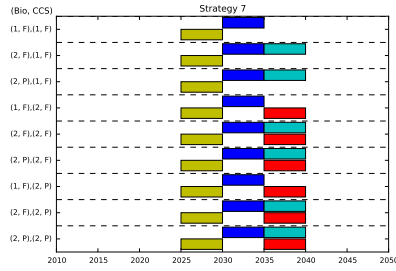


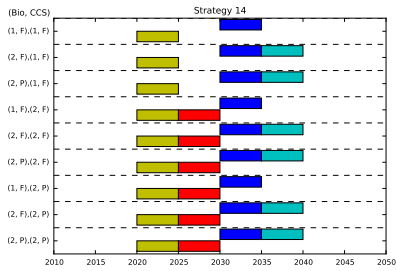
Figure 32: The pattern of strategies under different combinations of uncertainty front-loading with partial-revenues. The numbers refer to the strategies pictured in Figures 33 and 34



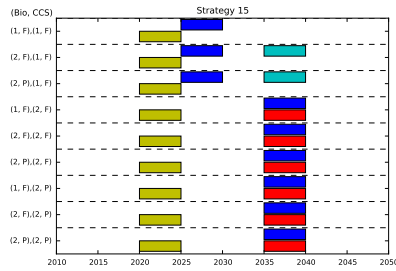
(a) Strategy 12



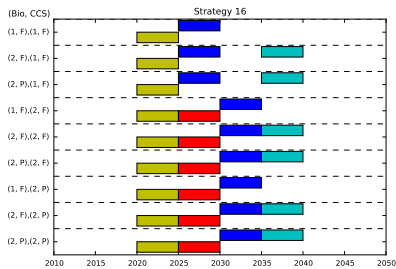
(b) Strategy 7



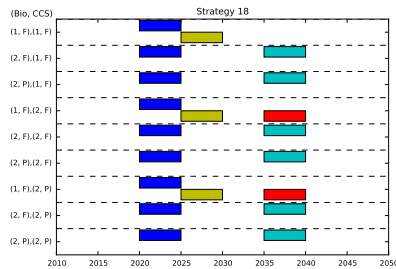
(c) Strategy 14



(d) Strategy 15



(e) Strategy 16



(f) Strategy 18

Figure 33: Six of the eleven main strategies from the example two-stage partial projects. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.

the situation when there are individual one-stage projects. The only difference is that the first-stage gives the option to continue to the second stage — an option to earn an extra revenue. This behaviour is a function of the split of revenues between the first and second stages of the research projects, which is derived from the ESME model results associated with the definition of the research project links to the ESME parameter values. In other words, even the availability of domestic biomass and build rates of CCS associated with the first stages of the research projects are invaluable. This value is manifest in how the research strategies alter when this is taken into account.

Strategy 18 in Figure 33f is a good example of an exploitative strategy. Investment in the first stage of the biomass project takes place in 2020. This leaves time to invest in the first stage of the CCS project before 2030 if the biomass project fails. In 2035, a second stage of investment takes place in the second phase of whichever project succeeded in the previous round of investment.

Strategy 7, shown in Figure 33b is a compromise between delaying investment, and recouping as much revenue as possible. Investment in the first stage of the CCS project occurs in 2025, recouping in 2030. The first stage of the biomass project following in 2030 no matter what the result from the previous investment round, giving the option to invest in the second stages of whichever project succeeds (or both) before 2040.

Focusing on the sequence of strategies along the y-axis of Figure 32a, we can see that below the threshold of around £1bn for the first stage of the CCS project (around £11 bn for the whole project), the chosen strategy is sensitive to the cost of the first stage of the biomass project. As the cost decreases below £13bn, strategy 8 gives way to 9 continuing through to strategy 12 below £7bn. These five strategies are displayed in Figure 34. This sequence displays a surprising progression. In Strategy 8, investment occurs only in the biomass project, despite the first phase of the CCS project being very cheap, and the biomass project being over £13bn. As the biomass project declines below £13bn, investment in the two phases of the CCS project occurs when the first phase of the biomass project fails (Strategy 9). This strategy matches Strategy 3 of the “big pharma” results (compare Figures 30d and 34b). Below £11bn, investment in the first stage of both projects in parallel in 2025 locks in the first-stage revenues in 2030. No investment in the second phases of the projects take place (Strategy 10). At £7bn, the second-phase investments are included in 2035 (Strategy 11, which matches Strategy 4 in the “big-pharma” results). Below £5bn, investment in biomass is displaced entirely by early investment in both phases of the CCS project for completion by 2030 (Strategy 12).

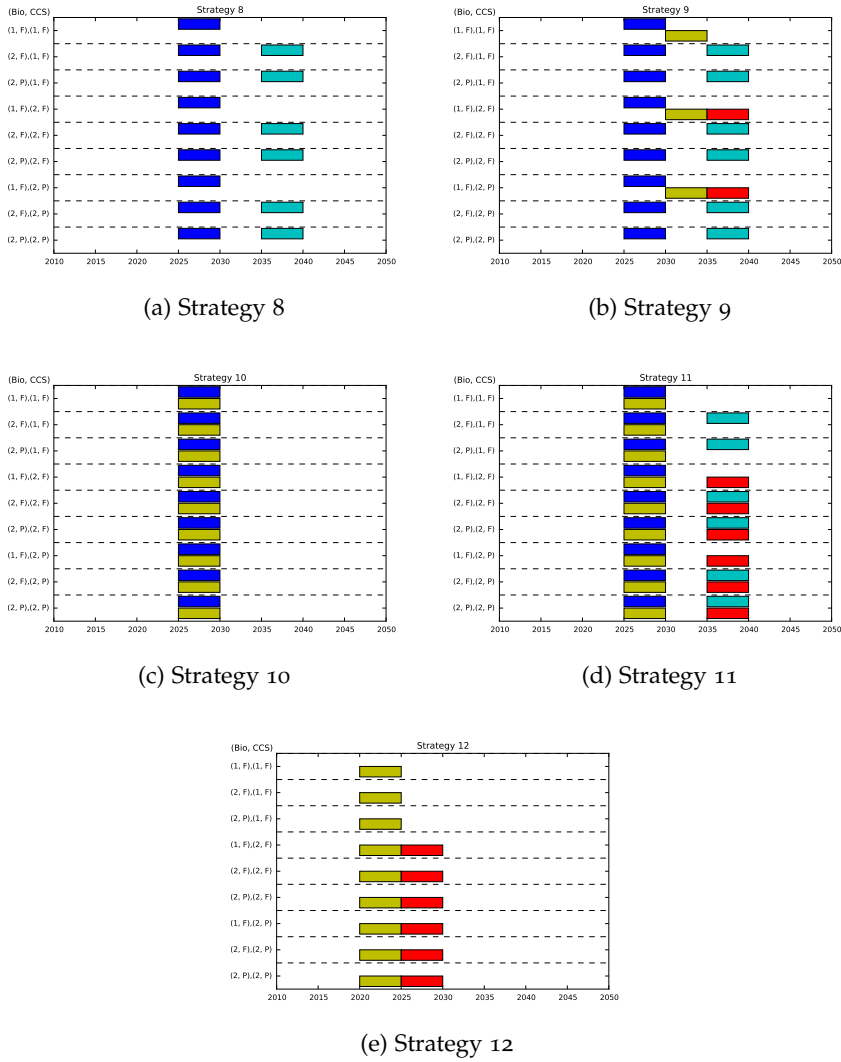


Figure 34: Five strategies from the 0.25/0.80 balance of probabilities with a two-stage project with interim revenues shown in Figure 32a. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.

6.4 ENERGY SYSTEM MODEL RESULTS

In this section, I outline the results from the [ETI-ESME](#) model for the UK energy system at different levels of availability of domestic biomass, and under differing build rates for carbon capture and sequestration plant. The combination of the two are then investigated. The results for the energy system model are presented using the structure of the [ESME](#) scenarios to build the decision tree associated with the potential success or failure of the corresponding [research project](#).

These results demonstrate the operational and structural implications for the energy system under the various investment strategies presented in the previous sections. These results are generated simultaneously with the total energy system cost which is used to populate the revenue function in the learning model.

To place the results of the learning model in context, I describe the general results from the energy system model, with a particular focus on how the changing availability of biomass and the build rate of carbon capture and storage technologies affect the wider energy system. Initially, all the other variables are fixed at their central values. For the majority of input parameters, this is unlikely to materially affect the results, as demonstrated by the findings from the sensitivity analysis. However, two of the top four most influential variables, liquid fuel price and natural gas prices, are also fixed. The sensitivity analysis suggests that the effects of changing these variables are linear, thus little interaction effects are evident. However, the potential strong (and important) assumption of holding these fossil fuel prices constant is not investigated and is left for future work.

Biomass availability and CCS build rate were identified in the sensitivity analysis as critically important parameters. I have devised a visualisation based on the [scenario trees](#) to investigate these parameters independently, followed by the combination of the two. The [scenario trees](#) for biomass and CCS with one-stage only correspond to 10 ESME scenarios when the projects are conducted independently. Given the [NACs](#) used, the model cannot pre-plan before an uncertainty is resolved. Hence, the [scenario tree](#) also shows how the future branches out from an initial condition, albeit in dramatically simplified form.

The [scenario tree](#) corresponds to whether domestic biomass becomes available (or CCS build rates increase) in 2020, 2030, or 2040. By 'become available' I mean as an [outcome](#) of a learning process, which as outlined in Chapter 5 is modelled as an irreversible [research project](#) of sequential [stages](#). If domestic biomass 'becomes available' in 2020, that is because there has been some research, prototyping process, supply-chain innovation or [commercialisation project](#) which has effectively made lots of domestic biomass available for use in the energy system. The same stands for the CCS project [scenario tree](#).

Once the biomass is made available through the success of the underlying [research project](#), it is available for the remainder of the model horizon.

The [scenario trees](#) on pages 156 to 162 show several different aspects of the ESME results. Firstly, the plot for the year 2010 shows a summary of the base year data. All the plots for subsequent years show the difference from the previous time period (hence the delta sign in the x-axis label). The plots are organised into columns corresponding to the five-year model time periods. The [scenario trees](#) of results allow comparison across the model results. For example - if biomass availability is low, or medium or high in 2020, the differences between the energy systems can be observed. Finally, comparing between the biomass availability and CCS build rate cases show dramatically different changes in the energy system. It is therefore interesting to analyse how the combinations of biomass and CCS work together or against one another in terms of the system structure.

Table 30: Sectoral emissions in the base year (2010) of the ETI-ESME model

Category	Emissions (MtCO ₂ e)
Buildings Sector	104.9
Industry Sector	59.5
Int'l Aviation & Shipping	40.2
Power Sector	178.8
Process and other CO ₂	40.4
Transport Sector	120.6
Total	544.3

Before delving into the parameter specific results, I explore attributes of the results common to all the cases. Several high-level patterns are evident across the cases. These are largely a response to the main emission constraint which reflects the ambition to reduce UK carbon dioxide emissions by 80% below 1990 levels. In loose agreement with many of the previous studies on this subject (e.g. Usher et al., 2010a), the ETI-ESME model indicates that the majority of effort occurs in the electricity generation sector. This is perhaps obvious, as in 2010, the base year from which the model begins optimising the future energy system, electricity generation comprises almost 33% of total emissions (see Table 30), more than any other sector. Under an optimistic case, where biomass is plentiful and CCS technologies are available, primary fossil resource use more than halves between 2010 and 2050, from 85% of the 2500 TWh in 2010 to 35% of the 2200

TWh in 2050⁴. Using an equivalent accounting method to the way that nuclear and fossil primary resources are calculated, renewable electricity and biomass resources grow from a 3% share in 2010, to over 30% in 2050. The exact shares change across cases, but the broad movement from fossil fuels to renewable and nuclear resources holds across all the cases.

As identified in the sensitivity analysis in Chapter 4, the availability of domestic biomass influences the total cost of the energy system. The importance of biomass is evident in how the results change across the ESME scenarios. Typically, the constraints on the availability of domestic biomass are mitigated somewhat by the use of more expensive imported biomass resources. However, this does shift the supply curve and results in new combinations of technologies across the sectors of the energy system.

6.4.1 *Availability of Domestic Biomass - Two stage*

The results from the energy system model are presented following the scenario tree for biomass availability used in the case study as shown in Figure 35. The use of three key-years, 2020, 2030 and 2040, and three stages of biomass availability, corresponding to 0, 90 and 180 TWh per year ([L]ow, [C]entral or [H]igh), result in ten scenarios that need to run through the ESME model. Biomass availability is divided into three levels, corresponding to the two-stages, plus a level corresponding to inaction. Due to the irreversibility of the research project, if biomass availability is [C]entral or [L]ow in an earlier period, then there is again an opportunity for it to be either [H]igh or [C]entral, or [H, C or L] respectively in the subsequent years. The reference case, against which the other scenarios are compared, is the situation in which no improvement in biomass availability takes place, thus domestic biomass is constrained to the minimum likely amount of 0 TWh per year. As visualised in Figure 35, all scenarios are identical up to the year 2020 when the uncertainty of biomass availability is observed to be either the same, or equivalent to the outcome of the successful completion of one or two stages of the research project. Numbering the branches of the decision tree in Figure 35 from bottom to top, scenarios 1 to 3 are identical to 2030, where the outcome of the research project (if any) results in the biomass availability associated with stages one or two of the research project. Scenarios 7 to 9 include a middling quantity of biomass availability, a value associated with the successful outcome of the first stage of the research project in 2020, followed by no change, or an increase in biomass availability to that associated with the successful outcome of stages one (C) or two

⁴ Note that this includes primary nuclear energy which is accounted for as for fossil fuels, while primary renewable energy, such as 'wind' is not accounted for in the same way - instead on a 1-to-1 basis

(H). Scenario 10 represents that if in 2020 the biomass [research project](#) is fully and successfully completed, thus the maximum quantity of biomass is available.

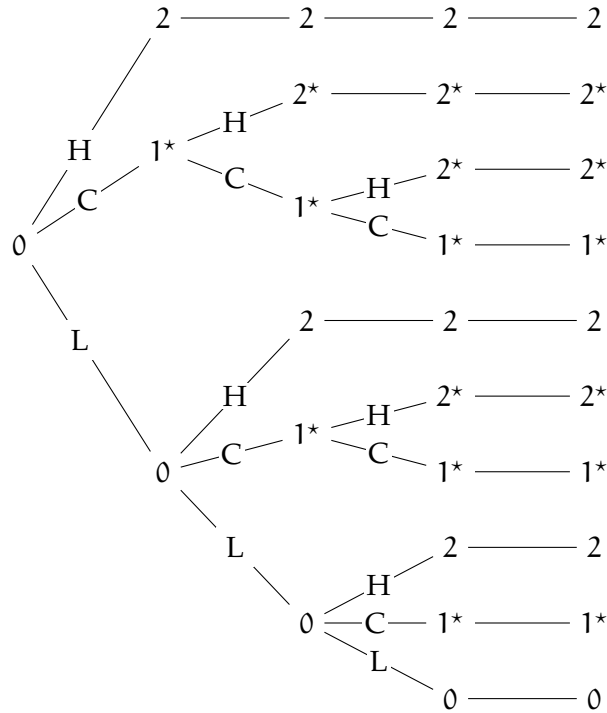


Figure 35: The scenario tree for biomass availability (two-stages, three key-years). The tree for the one-stage research project with three key-years is a sub-tree of this one, removing those nodes marked with a *.

Figures 36, 37 and 38 show the changes across the energy system according to different timings of biomass availability. The figures correspond to results for the change in installed capacity in the electricity sector, the changes in resource use, and the change in net CO₂ emissions.

The model has a strong tendency to delay installation of electricity capacity to the 2050 period in the biomass case. This is most apparent in Figure 36, where the 2050 plots are very similar in H/C/L cases, and the timing of when H/C/L is of negligible importance.

In Figure 37, the change in primary resource use occurs considerably earlier, with the available domestic biomass displacing coal (in all cases), natural gas and liquid fuel as biomass availability increases in 2030. If available, excess biomass in 2030 and 2040 is directed towards use in the biomass boilers in buildings. Beyond 2030, the changes in resource use are contingent upon the out-turn in the previous time period. But generally, in 2040, domestic biomass displaces biomass imports and liquid fuel which is used mostly in the transport sector. By 2050, the model changes are similar across the scenarios as the model struggles to meet the 80% CO₂ reduction tar-

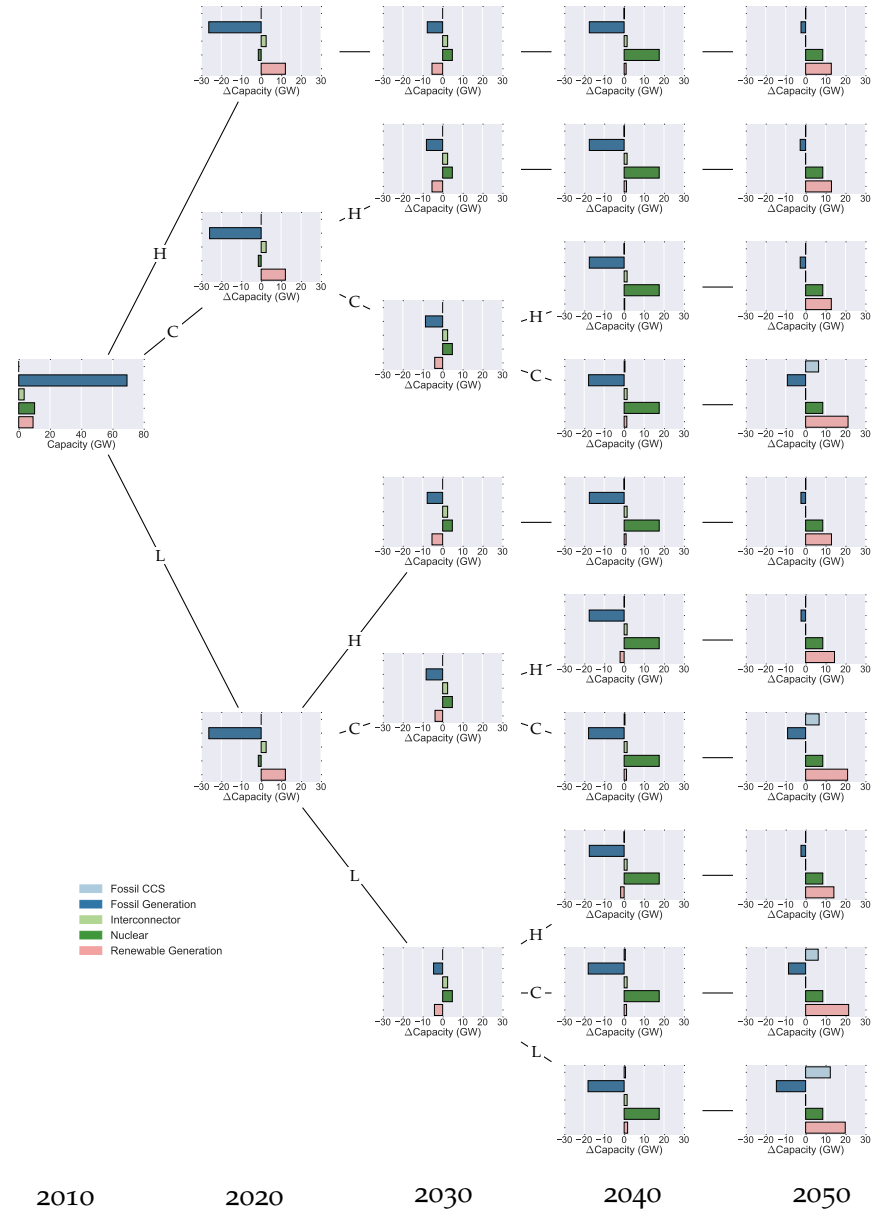


Figure 36: The change in installed capacity (GW) over time as a consequence of biomass availability

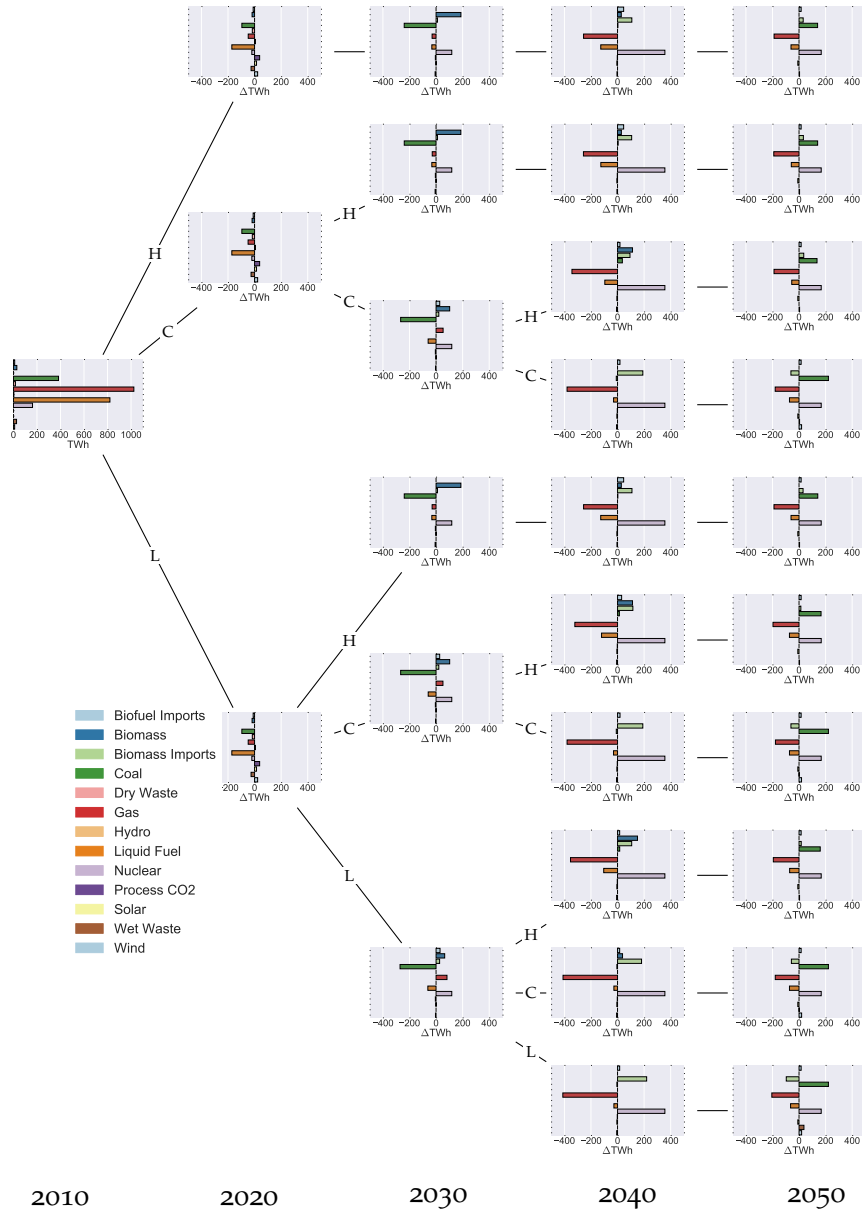


Figure 37: The change in resource use (TWh) over time as a consequence of biomass availability

get with a rebound in the use of coal (with CCS) and nuclear fuel. Coal use is inversely proportional to biomass availability in 2050, and reflects use in the small amount of CCS capacity that may be installed in the biomass scenarios.

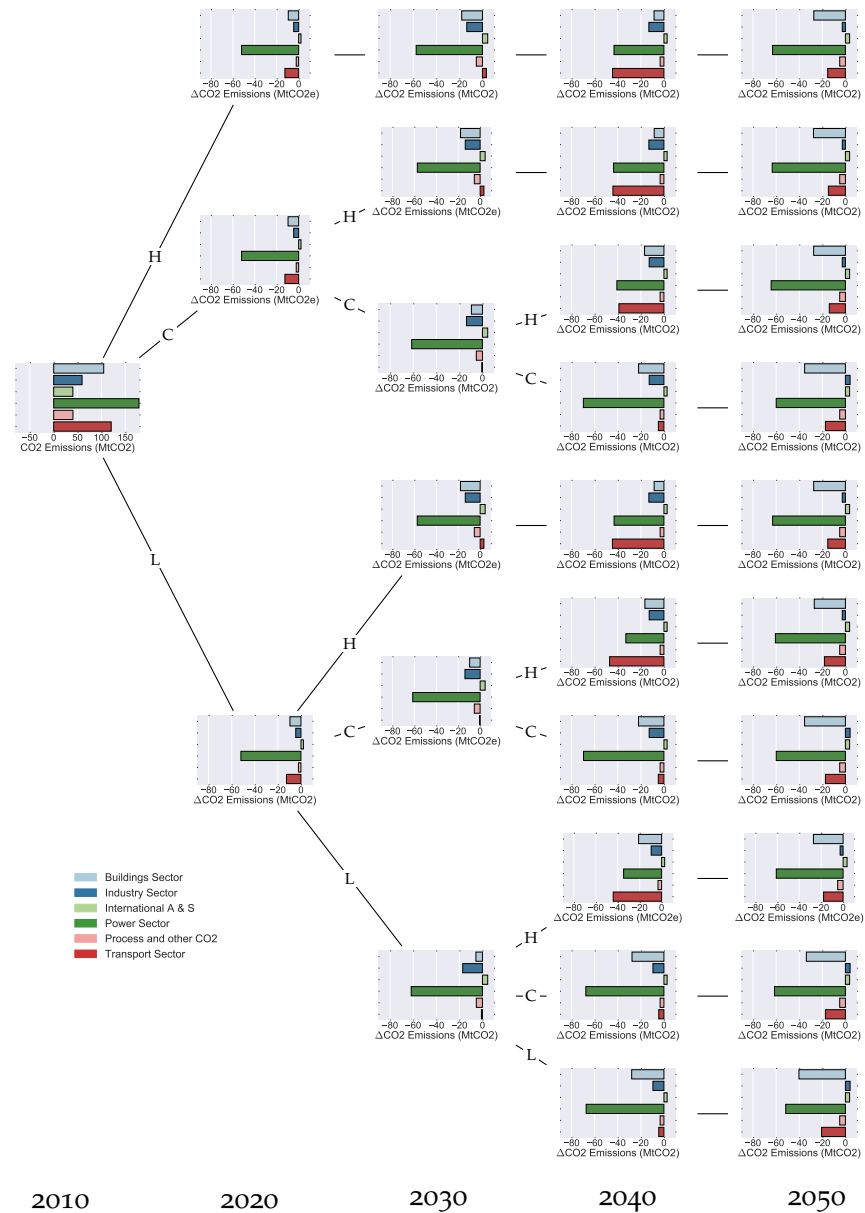


Figure 38: The change in net CO₂ emissions (MtCO₂) over time as a consequence of biomass availability

Figure 38 extends the results to the change in sectoral CO₂ emissions. Again, little change between the scenarios is evident in 2020. In 2030, the out-turn does affect the pattern of emissions reduction, but little evidence of irreversibilities are apparent. This cannot be said of 2040, where the change in sectoral emissions differs according to the out-turn of the previous period, 2030. For example, high biomass availability in 2030 brings forward decarbonisation of the building

and industry sectors to this period, otherwise, deeper cuts are made in the later period. High biomass availability also allows significant emissions reductions in the transport sector, with fewer reductions in the power sector, when compared against medium and low biomass availability.

6.4.2 *Build rate of Combined Capture and Sequestration Plant - two stage*

Figures 39, 40 and 41 show the changes across the energy system according to different timings of CCS build rate. The figures correspond to results for the change in installed capacity in the electricity sector, the changes in resource use, and the change in net CO₂ emissions.

As evident in Figure 39, the installed capacity of renewable technologies is inversely proportional to the availability of CCS, especially in 2020 and 2030. In 2030, under a high out-turn, fossil fuel generation capacity is reduced more than in the other cases in anticipation of installing CCS in 2040 and 2050.

In 40, the lack of difference between the scenarios is the most notable aspect. There are minor differences in the pattern of liquid fuel use in the earlier periods and a trade-off against biofuel imports. In 2050, there is a clear link between the build rate of CCS and the use of coal in that period. Under a high-build rate of CCS and given the confines of the carbon emission constraint imposed upon the energy system model, only then may coal be used, displacing the consumption of renewable resources.

In Figure 41, CO₂ emissions are shown as the net total for each year in the time period. The changes are subtle as the small changes are only apparent in 2030 and 2050. The distribution of emissions across the sectors is almost identical in 2040 across all the scenarios. What is apparent, is that an early increase in the build rate of CCS technologies does slightly increase the negative emissions in the power sector in 2040 and 2050, easing the pressure on the transport and buildings sectors. However, there is little apparent benefit to an increase in the build rate of CCS before 2040.

6.4.3 *Combining Biomass Availability with CCS Build Rates - One stage*

The combination of both biomass availability and CCS build rates reveals the value of the two uncertainties both independently and together. The combination of the one-stage research projects are shown here; the results of the two-stage give 35 scenarios which would be too numerous to summarise effectively. The one-stage results give the headline results at the extremes of the values for the two-stage results.

Domestic biomass and CCS technologies interact both directly and indirectly within the ESME model. Biomass may be consumed by conversion technologies to produce electricity, domestic heating, in the

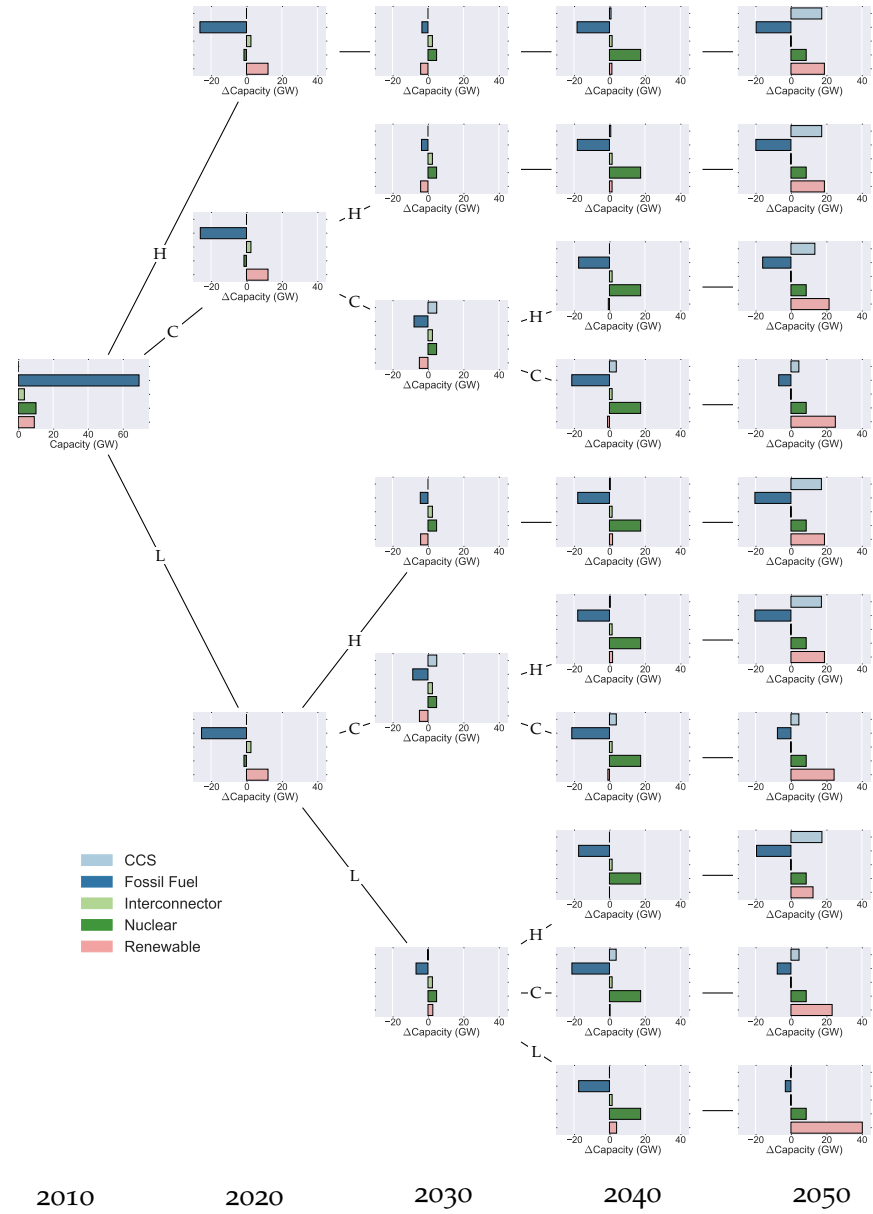


Figure 39: The change in installed capacity (GW) over time as a consequence of build rate of CCS

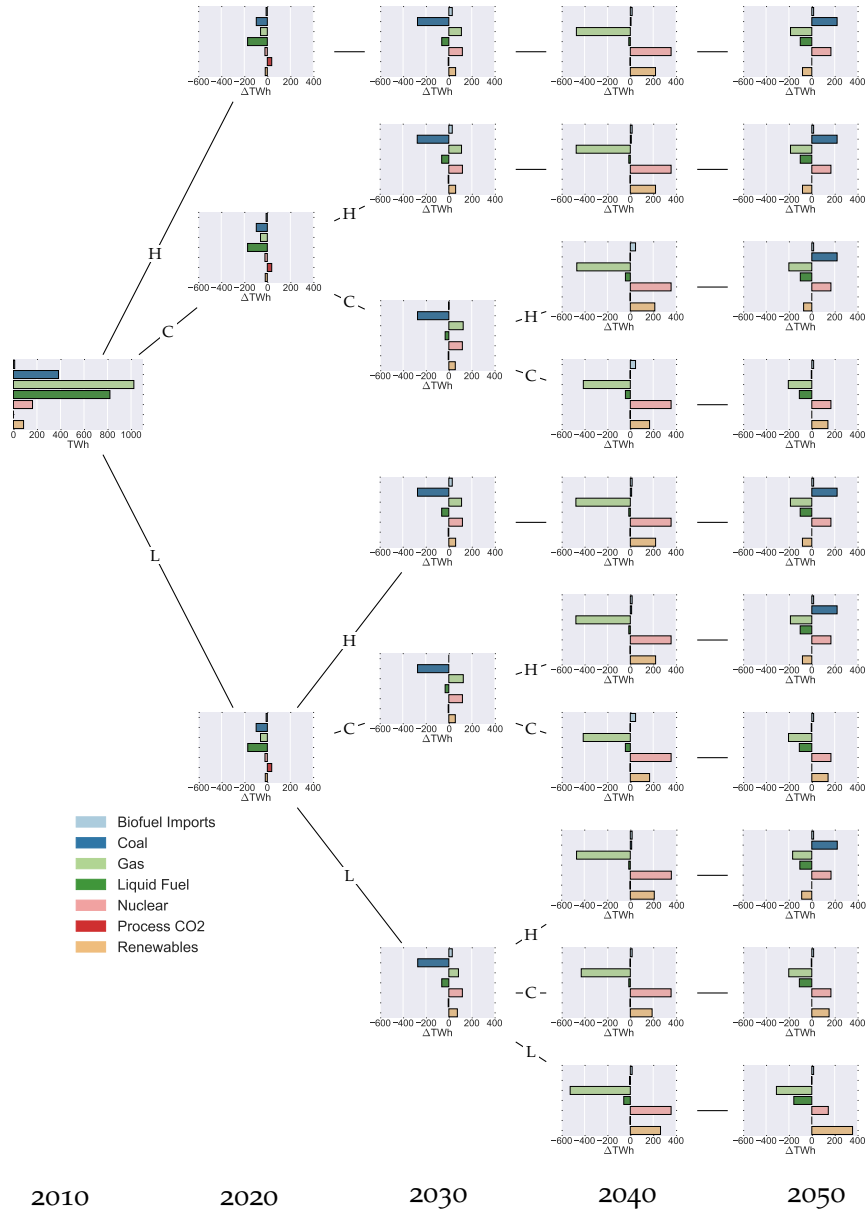


Figure 40: The change in resource use (TWh) over time as a consequence of build rate of CCS

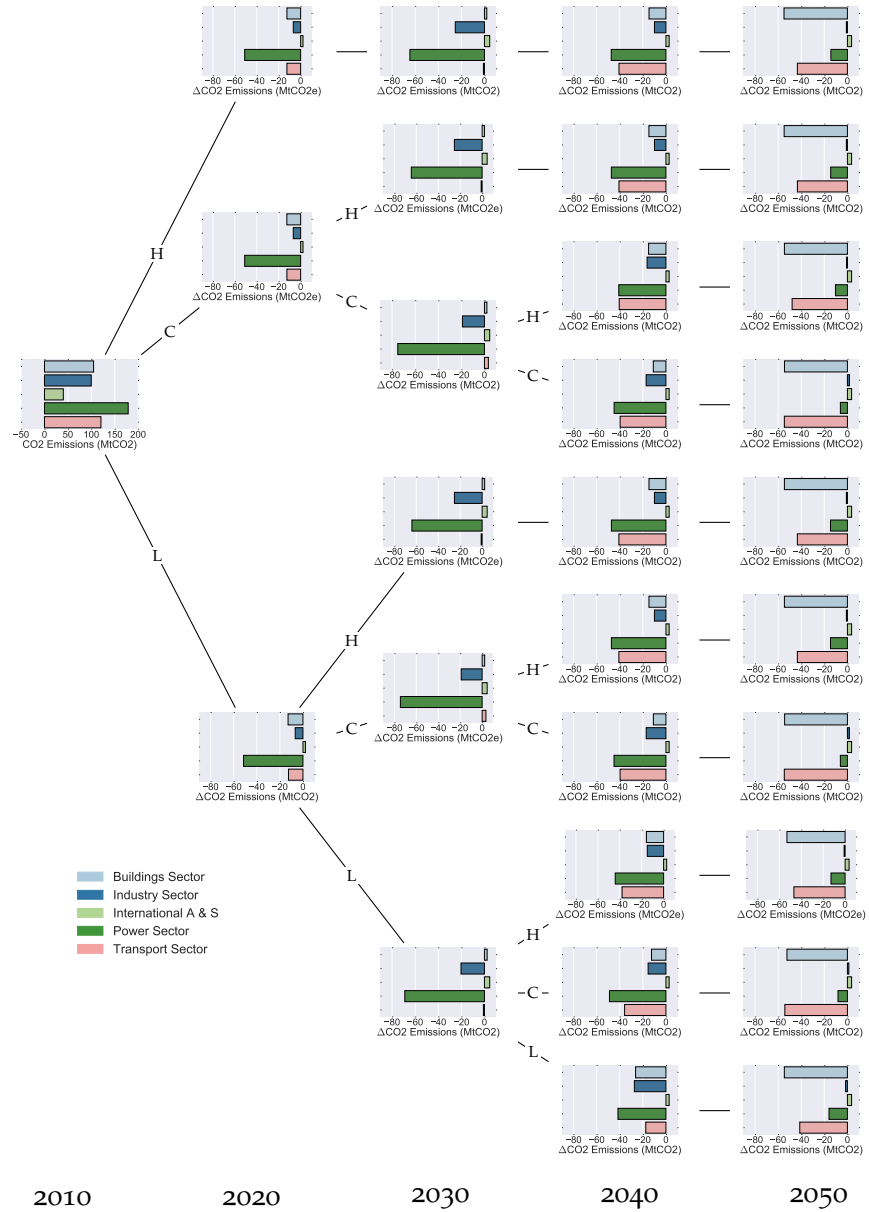


Figure 41: Annual net CO₂ emissions (MtCO₂) over time as a consequence of build rate of CCS

industry sector and can be converted into hydrogen or natural gas through gasification technologies. Biomass can also be used directly in CCS plants to produce ‘negative emissions’, or co-fired with coal in traditional or CCS-enabled plants. Finally, the gasification and industrial processes which consume biomass can also be sources for carbon sequestration. Certain conditions can make complex linkages of technologies, such as hydrogen production from biomass with CCS, a viable option, particularly as hydrogen can be used in typically difficult to decarbonise sectors, such as industry and transport.



Figure 42: Annual change in capacity over time as a consequence of build rate of CCS and biomass availability

Figure 42 shows how the availability of biomass reduces the capacity of renewable technologies from 2030 onwards. Having biomass means that less renewable capacity is needed in the electricity sector, but CCS does not have the same effect unless CCS is only available from 2040. Without CCS or biomass, a massive investment in renew-

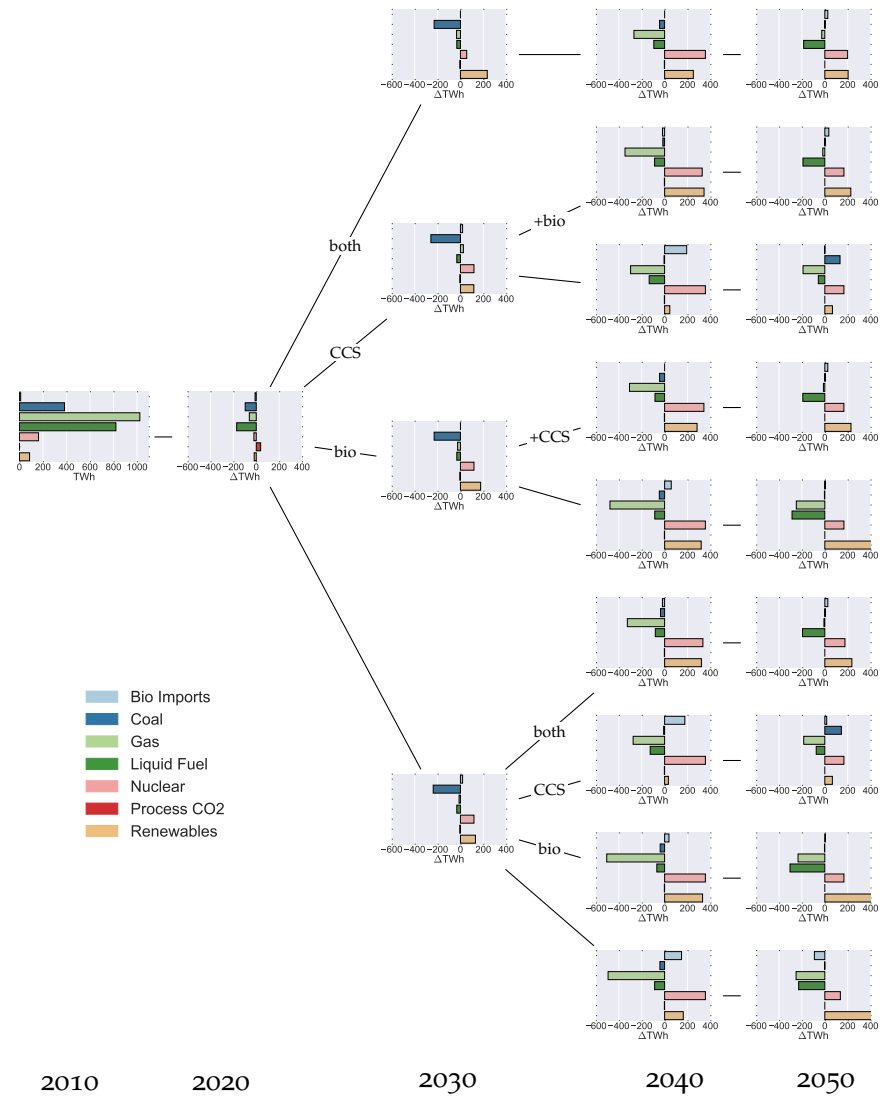


Figure 43: Annual change in resources over time as a consequence of build rate of CCS and biomass availability

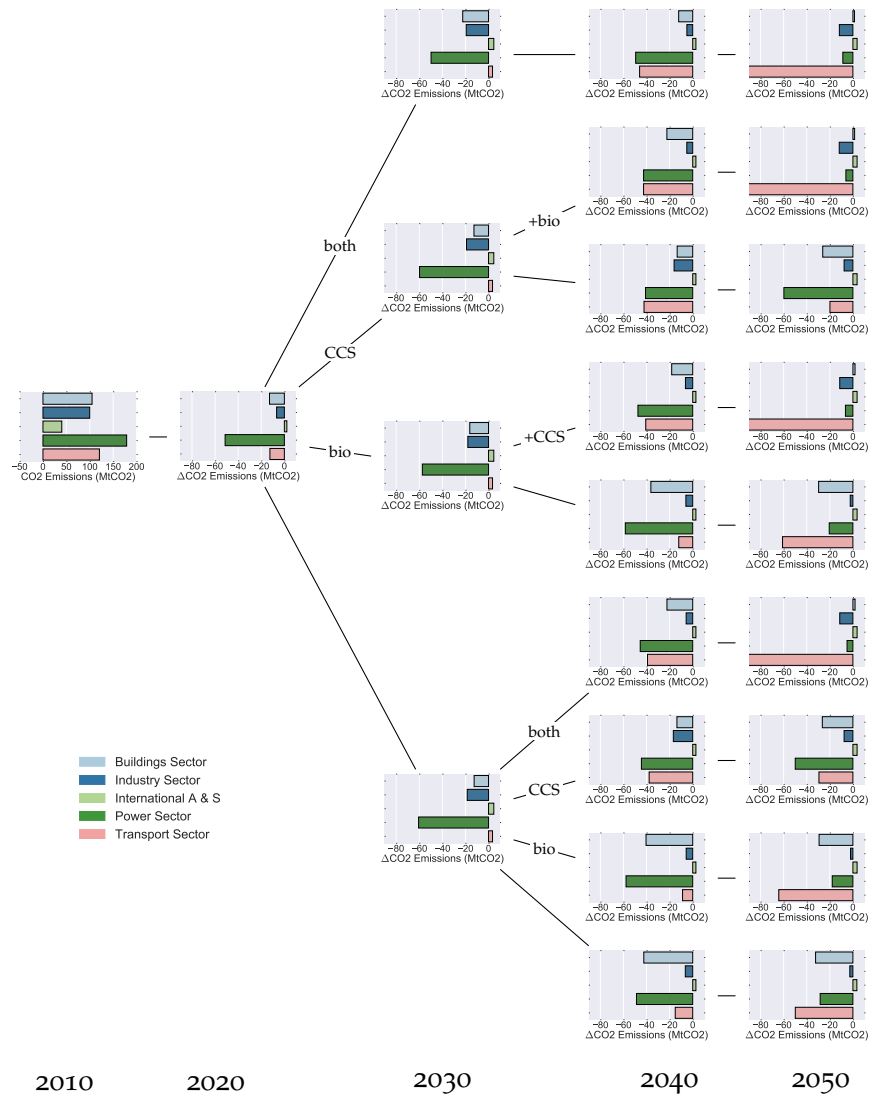


Figure 44: Annual change in net emissions over time as a consequence of build rate of CCS and biomass availability

able technologies is required in 2050 to meet the 80% reduction in CO₂ emissions. Late availability of biomass and CCS means that the energy system must catch up to achieve the optimal system structure for 2050. Nuclear capacity is unaffected by the availability of biomass or CCS.

Figure 43 shows how coal consumption is reduced across the system, particularly in 2030. As coal is the most carbon intense of all the fossil fuels, this is not surprising. However, the presence of CCS (but no biomass) means that coal is able to be used in 2050, the residual emissions acceptable due to the significant drop in emissions from the transport sector. In the absence of domestic biomass availability, some of the shortfall is made up from more expensive imports of biomass, but only when CCS is available. Biomass availability reduces the consumption of liquid fuel, particularly so in the absence of CCS.

In Figure 44, the availability of both biomass and CCS means that a large decarbonisation of the transport sector can take place, dominating the reduction in emissions of all other sectors in 2050. An greater early reduction in building emissions is also possible over alternative outcomes. If the availability of biomass and CCS is delayed, or doesn't occur, then much more effort is required in builds and power to offset the lack of reduction in transport emissions.

In this Chapter, I have presented the results from the learning and the ESME models. The results from the learning model demonstrate the extraordinary increase in potential strategies as the the number of research projects and stages increases. With one project, such as for domestic biomass availability, of just one stage, there are only three strategies. Which strategy is chosen as optimal is a function of the cost of the project, the revenue derived from successfully completing the project, and the likelihood of this outcome. Under constant revenue and probability of success, as the cost of a research project increases, the optimal strategy moves from early investment to late investment to no investment. Introducing a second project, such as for CCS build rate, which interacts with the first through the revenue function, results in a more complex sequence of strategic investment. Now, as the cost of one project increases, the optimal behaviour changes from investment in both projects, to delaying the least valuable project, to abandoning the least valuable, to investing in neither project. As the probabilities of the project success change, the values of the project costs at which the optimal behaviour switches between investment strategies changes. These boundaries of project costs demonstrate non-linear and non-features as these probabilities change. However, the range of different investment strategies do not change.

When research projects are divided into two [stages](#), the complexity of the learning model results increase. The results show how the division of projects into two [stages](#) results in an increase in value over an inflexible one-[stage](#) project. This is shown by the increased threshold of project costs under which it is optimal to invest in the projects. This increase in value is solely due to the introduction of the flexibility to abandon a project at an early stage.

The results from the [ESME](#) model demonstrate the operational and structural implications for the energy system under the various investment strategies presented in the results of the learning model. The results show how CCS technologies and the availability of biomass interact both directly and indirectly in decarbonisation pathways out to 2050.

DISCUSSION

In this chapter, I discuss the results in a broader context and present conclusions to the thesis. Earlier, I introduced the topics of uncertainty and energy system modelling. While uncertainty has received plenty of attention in the energy system modelling literature, previous studies have seldom investigated decision-dependent uncertainty in this context.

Investigating decision-dependent uncertainties places the decision maker in a position to influence learning through an investment of time or money (or both). The representation of this type of uncertainty has an important role in the modelling of learning processes, some of which could influence the cost of transitioning to a low-carbon energy system.

To identify potential learning projects, it was first necessary to determine which uncertainties in the energy system are in fact decision-dependent. And if so, whether resolving the uncertainty was likely to have any effect in the face of the challenging transition to a future low-carbon energy system. To do this, I identified a technique to systematically conduct a [global sensitivity analysis](#) on the [ETI-ESME](#) model. Due to the particular combination of computational demands and the nature of the data inputs used for the model (different natures of uncertainty¹) the Method of Morris was identified as a good compromise between useful added information, computational burden and the risk of ‘overfitting’ the inputs. The results from this process identified that a tiny subset of the inputs were critically important. And only a subset of these influential inputs were amenable to learning. The inputs identified were the Availability of Domestic Biomass and the Build Rate of Carbon Capture and Storage (CCS). These are notably different to the common focus on technology costs found in many previous studies which investigate learning.

Consequently, two learning projects (themselves uncertain) associated with the availability of biomass and the build rate of CCS inputs to the [ESME](#) model were used to explore strategies, probabilities and cost thresholds below which investment in reducing these uncertainties is profitable. The structure of the [research projects](#) — the number of [stages](#), the timing of revenues, and the probabilities of success — were found to be very important in determining this cost threshold. Even with two projects, the degrees of freedom available to the decision maker are considerable, and the results indicate that numerical modelling the schedule of staged, consecutive learning projects

¹ see Appendix C

is essential to determine when and in which order projects should be attempted.

In Section 7.1, I answer the research questions posed at the end of the introduction on page 9. In Section 7.2, I look at the implications of this work for UK energy policy. Both the results from the learning model and the [global sensitivity analysis](#) have some potentially interesting insights to offer policy makers. In Section 7.3, I bring together the implications this work has for research in the field of energy system modelling and learning. Finally, in Section 7.4, I address the limitations of this study, including the choice of methods, and how the chosen methods relate to the model, the limitations of the [ESME](#) model and other energy system models and the approach taken to address dynamic uncertainty. I conclude with a list of concrete ideas and suggestions for further work in this field.

7.1 ANSWERING THE RESEARCH QUESTIONS

In this section, I address the research questions posed in Section 1.6.1.

7.1.1 What is the most appropriate global sensitivity analysis method to identify and rank uncertainties in energy system models, given the nature, source and type of parametric uncertainties?

In Section 2.2 I outlined the available [global sensitivity analysis](#) techniques, and discussed the importance of global versus local methods. In Chapter 3 I identified the most appropriate method for identifying and ranking critical uncertainties in the energy system. The choice of the Method of Morris was dictated by several particularities of [ESOMs](#). Firstly, as outlined in Appendix C, there are a wide variety of types of uncertainty. Relatively few of the uncertain inputs to an [ESOM](#) can be quantified using probability distributions and therefore managed with a risk based approach. For some inputs we may find it difficult to provide anything other than a range, and for a few inputs, even that may prove difficult. Using a [global sensitivity analysis](#) which relies upon probability distributions, or relies too heavily upon a detailed portrayal of the input data would be counter-productive. In the context of this thesis, the intention of running a [global sensitivity analysis](#) is to determine which inputs are important despite the fact we know little about their values. The Method of Morris requires the definition of parameter ranges and samples from just four points within this range. This means that unless the distribution of an input is radically skewed, the assumption of uniformity is reasonable and does not bias the results from the sensitivity analysis. The second criteria was the computational demands of the [ESME](#) model, which run into tens of minutes per model run. The Method of Morris is computationally efficient requiring in the order of $10(k + 1)$ scenarios, where

k is the number of input parameters (or groups of input parameters). As such, the analysis allowed 119 parameters, formed into 31 groups to be analysed in 320 scenarios. The choice of [global sensitivity analysis](#) technique is a compromise. However, even given much larger computing resources, alternative methods such as the variance based ones of Sobol' and Saltelli would not improve upon the results of the Method of Morris, given the data constraints, even though the outputs are more detailed. The outputs from a more comprehensive sensitivity analysis technique would not be more convincing if they relied upon spurious input data. Such an approach would undermine the reason for conducting a [global sensitivity analysis](#) in the first place; transparency and information about the model behaviour over a large expanse of the model input space.

7.1.2 *What is the value of learning about critical energy system uncertainties?*

By running the [ESME](#) model over controlled permutations of the critical uncertainties, I computed the value of resolving these uncertainties over two key-years, 2030 and 2040, shown in Sections [6.1.3](#) and [6.2](#). The value was computed as a discrete look-up table of discounted total energy system costs, a [revenue function](#). This was generated by subtracting from the total cost of a reference scenario, in which no action was taken, the cost of a scenario in which an uncertainty was resolved in a key year and to a predefined level. By implementing [NACs](#) into the deterministic version of the [ESME](#) model, I was able to avoid the model anticipating the resolution of an uncertainty prematurely. For the two critical uncertainties investigated, the [revenue functions](#) ranged between £0bn and £72bn discounted to 2010 for the combination, up to £36bn for the availability of domestic biomass at 180 TWh/year and up to £43bn for a 2GW/year build rate of CCS plant. These values are relative to the results of the [ESME](#) model and were computed with two key exogenous uncertainties, the price of liquid fuel and natural gas resources, held fixed to their central values. These values represent a range of between 4% to \approx 17% of the discounted total energy system cost.

7.1.3 *What is the structure of dynamic energy-system uncertainties?*

Only two of the top five uncertainties identified in the sensitivity analysis were amenable to the learning model developed in Chapter [5](#). A distinction was made between exogenous and endogenous uncertainties. The former are external to both the model and the influence of the decision maker, and must therefore be managed using a risk-based approach. The latter are also known as dynamic or decision-dependent uncertainties, those which are subject to influence from

the decisions made in the model. Mapping the ranking of input parameters to the energy system model to these categories revealed that liquid fuel and natural gas resource prices are exogenous, while biomass availability and CCS build rate are endogenous.

7.1.4 *How do dynamic energy-system uncertainties relate to opportunities for learning or reducing these uncertainties?*

The results presented in Chapter 6 indicate that the structure of a [research project](#) has significant effect upon the profitability of the project. Firstly, dividing a project into two [stages](#) doubled the cost threshold under which investment would take place. Ensuring that the balance of the project [stage](#) probabilities were front-loaded (least likely to succeed first) increased the value of the projects over that of projects which delayed the uncertainty to the later [stage](#). However, when revenues are received for successfully completing interim [stages](#) of a multi-[stage research project](#), the insights become less clear-cut and case specific. For example in a two-[stage](#) project, as the revenue from successful completing the first [stage](#) increases to cover the cost of the second [stage](#), the balance of probabilities between first and second [stages](#) are unimportant. Instead, the cost of conducting the first [stage](#), and the probability weighted revenue from completing the second [stage](#) are all that are required to compute the net present value of the project. In other words, the project collapses back down to a one-[stage](#) project as the revenue from the first [stage](#) and cost of the second [stage](#) cancel one another out. This is explained in Section 5.1.2.

The optimally structured [research project](#) places the cheap, but unlikely [stages](#) first, delaying the more expensive, but more certain parts of the project, to later [stages](#).

7.2 IMPLICATIONS FOR POLICY

In this section, I discuss the implications for policy of this work. I first summarise the findings from the application of the learning model to the two dynamic uncertainties identified from the [global sensitivity analysis](#). I briefly compare the findings from the learning model with some of the figures from the energy research literature pertaining to the UK. I conclude with a summary of the policy insights from the [global sensitivity analysis](#).

7.2.1 *Results from the Learning Model*

The main implications for policy from the learning model are:

- Individually, the availability of 150TWh/year of domestic biomass is worth in the order of £36bn, while the ability to build CCS plant at a rate of 2GW/year is worth up to £43bn. Together, the value increases to a maximum of £72 bn
- The structure of the [research projects](#) were found to have a large effect upon the cost threshold below which investment is optimal
- When a [research project](#) was structured as a one-phase *moon shot*, and the probability of the project succeeding is 20%, investment in both projects proceeds when costs are below £10bn
- When split into two distinct [stages](#) but otherwise identical, the cost threshold below which the project investment is optimal doubles to around £22 bn
- Within these cost thresholds, there are a large number of different investment strategies, some of which are counter-intuitive
- These findings indicate that there is unlikely to be a substitute to the modelling of decision dependent uncertainty

I address these points in greater detail in the sections below.

7.2.1.1 *Biomass and CCS are worth billions of pounds to the UK energy system*

Results from the [ETI-ESME](#) model revealed that in the absence of these critically influential technology/resources, the costs of the energy system increase by tens of billions of pounds². Individually, the availability of 150TWh/year of domestic biomass is worth up to £36bn, while the ability to build CCS plant at a rate of 2GW/year is worth up to £43bn. Together, the value increases to a maximum of £72 bn. The timing of when biomass and CCS become available at scale is important—little benefit is experienced before 2030, but there is a large benefit to having successfully completed both projects by 2030 (£72 bn) rather than 2040 (£62 bn).

This study improves upon previous deterministic energy system model studies by incorporating some measure of the uncertainty associated with the development of novel technologies or systems. Previous studies have not incorporated the non-anticipativity constraints used in this study and potentially under-estimate the costs of a transition to a low-carbon energy system. For example, included in the revenue function is the cost of the *surprise* of finding out whether the technology or resource actually is available in a specific year. In contrast, a deterministic, perfect foresight model has perfect knowledge of when the surprise occurs and is therefore able to plan accordingly.

² All costs are discounted back to 2010

This both underestimates the true costs of having the technology or resource, and underestimates the effect of timing and path-dependency. Myopic approaches (as pointed out by Anadon et al., 2011, p. 106), in which the time horizon over which investment decisions are made is divided into chunks, would go some way to replicating the effect, but only by ensuring that previous decisions are locked in place. The non-anticipativity approach does this in a more comprehensive way, by exploring the branches of the [scenario tree](#) associated with the degrees of freedom afforded to the decision maker, within the constraints of the [stages](#) defined.

Taken together, the value of the technology/resource pair is not quite the sum of the two individual projects. This illustrates the importance of taking the options into account together—the combined value of the options is non-linear. This is likely to be exaggerated as a larger number of options are integrated into the analysis. Some of the mitigation options evaluated in the individual cases may be mutually exclusive, and mitigation options which alone can have a large effect, may have much less of an impact when combined with ‘competing’ mitigation options.

7.2.1.2 *The structure of research projects modifies the cost threshold below which investment is optimal*

The results compare single-[stage](#) and two-[stage](#) projects with otherwise identical outcomes. The division of the project into [stages](#) increases the flexibility to abandon the project, increasing the likelihood of investment in the project. However, the timing of revenues, and the probabilities that define the successful completion of the [stages](#) alter the optimal investment strategy. With one-[stage](#) projects the only flexibility the learning model has, is to determine the order in which investments should take place (including in parallel), and the choice of the projects. With two-[stage](#) projects, the number of decision points more than doubles, as the timing of each of the [stages](#) and the decision to abandon or continue a project under certain conditions becomes an option. The results show how increasingly complex behaviours and strategies arise as a consequence of allowing the learning model great degrees of freedom. These include hedging, exploitative and adaptive strategies.

7.2.1.3 *Investment in the one-stage research projects proceed when costs are below £10bn and the probability of the projects succeeding is 20%.*

With one-stage only, the investments follow a straightforward logic, with investment delayed as the cost of the project increases. If the costs of both projects are below £10 bn, then investment proceeds into both projects, although the exact timing and order of the investments

is a complex function of the combination of costs. When the cost of one project exceed the £10bn threshold, investment switches to the other project. When the costs of both projects are above the threshold, investment does not proceed and it is more cost effective to take no action.

7.2.1.4 *When split into two distinct stages but otherwise identical, the cost threshold below which the project investment is optimal doubles to around £22 bn*

Dividing a [research project](#) into two-or-more consecutive [stages](#), gives the learning model much more flexibility in the timing of investments, and the choice of combinations of project [stages](#). An immediate result of this is that the cost threshold under which projects are viable more than doubles, despite the fact that all other aspects of the projects remain equal. Part of this increase in the cost threshold is because the discounting of the costs takes place over two five-year periods, rather than one five-year period. However, the simple act of dividing a project into two [stages](#) increases the value of the project due to the flexibility introduced. Note that this result is valid under the no-regret style objective function, where the value of the research pipeline is maximised, with no funding constraints.

7.2.1.5 *There are a large number of investment strategies*

Optimal investment strategies are a function of the number of projects, [stages](#), probabilities associated with the success of the project [stages](#) and costs of the [stages](#). Without explicitly modelling the projects on a case-by-case basis, it is difficult to identify the optimal portfolio and timing of investments.

7.2.1.6 *These findings indicate that there is unlikely to be a substitute to the modelling of decision dependent uncertainty*

The relationship between the value of an uncertainty and the cost threshold of a [research project](#) is complicated by a number of factors. These include the interactions between the uncertainties influenced by the [research projects](#), the number of [research projects](#), the number of [stages](#) and the costs and probabilities associated with each of these [stages](#). The modelling approach used in this thesis enables opportunities to be identified to increase the value of a research portfolio through the strategic restructuring of [research projects](#), the selection of a research portfolio, and adoption of strategies which adapt to changing conditions.

The learning model enables analysis of the complex interactions between the timing of investment in uncertain research projects, and the interplay of technology investment in the energy system over time. For example, if a research project identifies that biomass is not

available (i. e. the project fails, perhaps contrary to expectation), then the learning model still identifies the optimal recourse strategy given this new information. Under certain conditions, the resolution of one research project (successful or otherwise) has a value for the other research project, through the interactions which are evident in the investments in the energy system model. These insights are categorically different to those obtained from models which do not investigate decision dependent uncertainty.

7.2.2 Using ESOMs to Inform the Energy R&D Agenda

In the introduction, I observed that that **ESOMs** are rarely used in the UK at the level of decision making about research programmes. The predominant use has been to apply the models in a policy role, to direct and explore opportunities for market-based policy and other government led interventions. They are used at a broader level for informing strategic decisions into research direction, and with the **ESME** model, **ETI** have begun a positive trend to inform investment decisions through quantitative analysis of long-term energy system trends.

Baker et al., (2006, 2008) investigate the effect of investment in energy R&D projects at a high-level by modifying a marginal abatement cost curve in an **IAM**. By considering the climate damage conditions under which both risky and incremental projects are adopted, the results are ambiguous. The results from this thesis are rather less ambiguous, namely as there is no exogenous uncertainty represented in the carbon constraint in the **ESME** model (essentially a proxy for climate damages at the global level³).

Anadon et al., (2011) and Anadon et al., (2014) examine the state of energy R&D funding in the USA, using expert elicitation to estimate the technological cost reductions, and modelling the effect of those cost reductions using an energy system model. This work optimizes R&D funding for each area for different energy system indicators (such as CO₂ price, consumer and producer surplus, CO₂ emissions) using latin-hypercube sampling with no assumption of independence across 25 technology areas. In contrast, this thesis computes the value of critical energy system uncertainties using an energy system model, identifies which are amenable to R&D and then computes the optimal cost threshold of a related R&D project, given the likelihood of the project succeeding, and the interaction of the uncertainty with others in the energy system. The learning model takes into account the temporal trade-offs, the ability to adapt to failed projects, and selects a robust portfolio of **research projects** under uncertainty.

³ Actually, the UK's legally binding abatement target of an 80% reduction in greenhouse gas emissions below 1990 levels already take the uncertainty surrounding the level of damage due to global emissions

It is interesting to compare the cost thresholds from the results of the learning model with the existing expenditure on UK energy research and development. For example Skea, (2014) examines the R&D spend of IEA countries, which is measured in the tens of billions of US\$. The peak values are in the order of ²⁰¹²£14bn per year. In this study, improvements in biomass availability and CCS are individually worth a one-off sum of ²⁰¹⁰£30-34bn to the UK energy system alone, and almost double this together. As a sense check of investment required versus expenditure, this comparison shows that the values are of the right magnitude. However, UK expenditure on energy R&D is a considerably smaller chunk than the tens of billions of dollars reported across all the IEA countries. For example, Winskel et al., (2014) discusses how the recent surge in UK public-funded energy R&D peaked around 2010 with investments of ²⁰¹²£0.5bn across all energy sectors, including energy efficiency, nuclear, renewables and fossil fuels.

The learning model shows how the cost threshold of individual learning projects associated with these uncertainties increases above ²⁰¹⁰£20bn under structured portfolios of projects with likelihoods of success around 20%. As the likelihood of success of an individual project decreases, the cost threshold also decreases. Across the portfolio of projects, the interplay of probabilities and project costs is difficult to unpick without using the learning model.

The cost thresholds are a function of a collection of the assumptions contained within the [ETI-ESME](#) model, which generates the [revenue function](#), as well as those surrounding the likelihood of the success of a [research project](#). It is therefore difficult to make a concrete comparison to the existing level of funding, without understanding the decision makers' beliefs of the research portfolio.

Anadon, (2012) points out that non-R&D funding, including market based measures such as renewable portfolio standards, can achieve the same effect as direct-funding. While the UK directs less than 1% of its GDP towards R&D, there has been a comprehensive programme of market-based measures, such as the Renewables Obligation, which provided financial incentives for the commercial installation and operation of some renewable technologies. Feed-in tariffs have also been used to support fledgling solar photovoltaic and other technologies, although these support mechanisms are being incrementally withdrawn.

The definition of a [research project](#) in the learning model has been left deliberately vague and could be applied to both public-funded R&D and a [commercialisation](#) programme. A multi-stage [research project](#) in the learning model could well be used to represent this sequential process of moving from lab-based research to full commercialisation.

7.2.3 *Policy implications of the Global Sensitivity Analysis*

The main implications for policy from the sensitivity analysis are as follows:

- The critical uncertainties are split between exogenous and endogenous types, so that only a subset are amenable to research and development
- The availability of domestic biomass and the build rate of CCS are the most influential uncertainties

I address each of these points in more detail.

7.2.3.1 *Critical uncertainties are split between exogenous and endogenous types*

The results show that both exogenous and endogenous uncertainties are among the top five most influential inputs to the [ETI-ESME](#) model. The uncertainties require very different management strategies. The former are not amenable to investment actions by a decision maker but must be managed defensively through hedging. The latter are amenable to research or an equivalent decision-based approach. This implies that decision dependent techniques, and hedging or risk-based approaches should be used in conjunction with one another.

7.2.3.2 *The availability of domestic biomass and the build rate of CCS are the most influential uncertainties*

To reiterate the previous point, the availability of domestic biomass, and the build rate of CCS were identified across the outputs as the most influential input parameters to the model. Both are system-wide innovations which touch many sectors of the energy-economic system. Biomass can be consumed for domestic heating, in conjunction with CCS electricity generation plants to produce negative emissions, and it can be used as a feedstock for producing bio-fuels. Fundamentally, it is low-carbon if produced under *sustainable* criteria. Likewise, CCS is can be applied as a sticking plaster to multiple sectors, but particularly within the industrial sector - sequestering the emissions from non-energy related processes (such as cement production). However, CCS is also an enabler for seemingly exotic chains of energy processes - used in conjunction with biomass or fossil feed-stocks to produce hydrogen, which in turn may be used in the otherwise expensive-to-decarbonise transport sector.

7.3 IMPLICATIONS FOR RESEARCH

In the introduction, I suggested that the recognition of dynamic uncertainties, i. e. that the observation of uncertainty can be dependent

upon a decision to invest, has not been modelled in energy system models (with some notable exceptions, including Baker et al., 2015; Santen et al., 2016). Relaxing the assumption that learning is exogenous to the decisions, could highlight opportunities for investment that would otherwise be ignored. The assumption in existing energy system models which take uncertainty into account is that all uncertainties are exogenous. For example, when uncertainty is propagated through an ESOM using a Monte Carlo framework, this is exhibited in the outputs as a wider distribution of outputs, incorporating the assumption that learning takes place. In a stochastic programming framework, the model hedges against the uncertain values to mediate the effect on energy system cost. Within the parameter values explored, the decision to resolve an uncertainty is an opportunity with significant financial reward. This value is dependent upon the perceived likelihood of success of a research project. Even if the likelihood of success were very low, then the structuring of a research project into stages would allow an otherwise non-economic project to proceed.

7.3.1 Reflections on Alternative Methodologies

A systematic approach to uncertainty has been taken in this work to identify and quantify the influence of uncertain inputs upon the outputs of the ESME model. One issue that has been highlighted is the apparent contradiction between the nature of the uncertainty surrounding the inputs to ESOMs, the data needs of existing global sensitivity analysis methods, and the predominant technique used to manage these uncertainties within a consistent framework: scenario analysis.

As discussed in Appendix C, scenarios are an appropriate response to a particular type of uncertainty, when it is difficult to quantify the probability of an event, while the event itself is well defined. A good example is future UK economic growth, a key driver of growth in final energy demand. If we accept that scenarios are a methodologically sound way of managing uncertainty, how do we incorporate the findings from the sensitivity analysis performed in this thesis?

The global sensitivity analysis results in this thesis provides strong evidence for a systematic approach to analysing ESOMs using a global sensitivity analysis method to identify critical uncertainties. Once identified, these critical uncertainties can then form a core part of a scenario analysis using the ESOM. The benefits of such an approach are that the scenarios which are formed around the influential variables identified can be guaranteed to cover a significant proportion of the model's output space. This point that is especially salient given that relatively few inputs are likely to be responsible for the majority of variation in the model's output.

The alternatives look less appealing. Without conducting a [global sensitivity analysis](#), the alternative [scenario analysis](#) method is to identify the set of variables to vary through an expert model-user who is familiar with the behaviour of the model. Or, randomly choose inputs that are deemed interesting or important. In both of these alternatives, the likelihood of Type II errors⁴ are high. Another alternative is to use standard risk-based approaches, propagating uncertainty through a model using a Monte Carlo sample, while quantifying the uncertain inputs using expert elicitation (Usher et al., 2013). The disadvantages here, are that expert elicitation is time consuming, and the subjectivity of the results, though explicit, may not be a desirable quality. Ultimately, while [global sensitivity analysis](#) does provide a systematic method for assessing the importance of uncertain parameters, the results also rely upon the (also subjective) choice of bounds over which the analysis is run. So [global sensitivity analysis](#) and expert elicitation methods are complementary in that expert judgements can assist in uncertainty quantification, while global sensitivity analysis can highlight which uncertain parameters should undergo greater scrutiny.

However, [global sensitivity analysis](#) also has benefits for those who wish to cobble a risk-based approach onto an energy system model. For example, a factor fixing setting for the sensitivity analysis allows the dimensionality of a model to be reduced through identifying those inputs that can safely be fixed to their mean value. This allows the focus of a risk-based approach to shift to quantifying only the most influential parameters ensuring coverage of the majority of the model's output variation.

[Global sensitivity analysis](#) maintains a focus on parametric uncertainties, in lieu of structural uncertainties. However, by [parameterising](#) structural changes in a model, a [global sensitivity analysis](#) would still be useful in quantifying the relative importance of these structural and parametric uncertainties.

In summary a [global sensitivity analysis](#) is an essential tool to identify and rank the influence of model inputs. Typical scenario analyses aim to permute input parameters over wide ranges, to explore plausible, but dramatic differences in input values. Thus, exploring plausible ranges of inputs in a sensitivity analysis technique is likely to enhance the relevance of the results from a [scenario analysis](#).

7.3.2 *Model Insights from the Global Sensitivity Analysis*

The main implications for research from the results of the [global sensitivity analysis](#) are as follows:

⁴ Failing to incorporate an influential input into the [scenario analysis](#) resulting in a study which is biased, omitting a considerable proportion of the model output space

- Resource prices present a considerable source of uncertainty to the energy system model, particularly the prices of natural gas and liquid fuels
- A small number of uncertain inputs dominate the results of the [ESME](#) model
- No individual technology dominates any other in the results of the sensitivity analysis

I address each of these points in more detail.

7.3.2.1 *Resource prices present a considerable source of uncertainty to the energy system model*

Resource prices, particularly oil and natural gas prices, are volatile, with the past decade showcasing fluctuations in oil price between 2010\$20/bbl and 2010\$140/bbl. The finding that the prices of oil and natural gas are especially influential to the model results is worrying. As mentioned in above, these uncertainties are by their nature exogenous, and to a large extent aleatory, and so can only be managed through risk-based approaches, as far as probabilistic methods can be extrapolated to future values. The implications for energy system modellers are potentially important. Decision makers should consider whether a least-cost optimisation model is an appropriate tool if the results of the model are heavily influenced by the price of oil and natural gas. This consideration should take into account what results are of importance. For example the influence of oil and natural gas prices were shown to be linear in nature in the [global sensitivity analysis](#), and had considerable effect on the total system cost. However, the change in oil and gas prices did not affect the findings concerning which technologies are best performing (interaction effects were relatively small). Alternative objective function formulations, including robustness could be developed to focus on developing a technology portfolio that is robust whatever the oil and gas price, particularly as these prices are unknowable.

7.3.2.2 *A small number of uncertain inputs dominate the results of the model*

To experienced users of sensitivity analysis techniques, this is not a surprising finding. It is very unlikely that all the parameters in a model would influence the output to the same degree, and therefore there will be a concentration effect where a small number of parameters produce the most variation in the output. However, for the energy system modelling audience, this finding does give some pause for thought. The results show that this concentration is valid across all the outputs examined, including total cost, install capacity and the

percentage of renewable electricity capacity. If this is the case, why is so much energy and resource expended upon very detailed bottom-up representation of the energy system, when a relatively small proportion of the inputs to the model account for the greatest variation in the output? This analysis demonstrates how [global sensitivity analysis](#) is an invaluable tool to help direct analysis and validate energy system models. The outputs from the sensitivity analysis can also help direct the choice of inputs to include in a [scenario analysis](#).

7.3.2.3 *There is no individual dominant technology in the results of the sensitivity analysis*

To some extent, this is a function of how individual technologies can substitute for one another within the energy system model, and representative of the necessarily aggregate way in which the sensitivity analysis was performed, by grouping together similar technology groups⁵. However, despite these caveats, the results do present a ranking of technology themes or groupings above any individual technologies: presented in order of influence these are: the ability to sequester carbon through CCS, the coupling of CCS with biomass and availability of domestic biomass resources, nuclear and renewables as low-carbon vectors to generate electricity and efficient and alternative fuelled passenger transport.

7.3.2.4 *Discussion*

The sensitivity analysis raised a number of important methodological insights. [ESOMs](#) are known for their tendency toward ‘penny switching’ — rapid oscillation between competing technologies based upon minor changes in assumptions. The sensitivity analysis⁶ reveals this tendency by ranking the influence of *individual* technology costs low in comparison to other parameters, although aggregate groups, such as all the cumulative effect of all renewable technologies, do rank higher. This is because the [ESME](#) model has a large number of options for technology substitution, and in a model of perfect foresight, a more expensive technology will be substituted by a cheaper one. In essence, different combinations of costs for similar technologies merely rearrange the supply cost curve for different energy services. This may have a minor effect on the technology mix from scenario to scenario, but the total energy system cost is not affected.

⁵ Note that although the main results presented in the thesis used more aggregate groups of technologies, preliminary results which used different groupings, support the general finding that within each of the groups, no individual technology cost dominated

⁶ The results for technology cost may have been influenced by the choice of uniform distribution. There is a possibility that the choice of uniform distribution could place undue weight upon the tails for some competing parameters, depressing the prevalence of technology cost.

This playing-down of the influence of technology costs contradicts much of the work in the [ESOM](#) literature. The results from the [global sensitivity analysis](#) suggest that the focus should instead shift to key resource costs, including natural gas, liquid fuel (oil), the availability of domestic biomass, and the rate at which CCS technologies can be built. To some extent, these results are a function of the imposition of a fixed constraint upon carbon emissions. Under a much less stringent emission target, it is unlikely that biomass and CCS would be as important to the energy system as indicated by the main results from the [global sensitivity analysis](#). In the absence of an emissions constraint, as shown in Figure 10 on page 81, the most influential parameters remain fossil fuel prices, biomass cost and availability, and the cost of cars.

However the results of the [global sensitivity analysis](#) are also presented in terms of the influence upon the marginal price of carbon dioxide. In the absence of a regulatory cap on emissions, these latter sensitivity results show that the availability of biomass and CCS build rate would dramatically affect the marginal price of carbon which to achieve an equivalent reduction in emissions, either through a carbon tax, or a trading scheme. It must be remembered that these are the results from a single model. Repeating a similar [global sensitivity analysis](#) procedure on other models would increase confidence in these findings.

7.4 LIMITATIONS AND FUTURE WORK

In this section, I address the limitations of this thesis, discuss the implications of this work for future research and suggest some logical next steps. The results of the [global sensitivity analysis](#) presented in Chapter 4 are discussed in the context of the methodology shown in Chapter 3. In Section 7.4.5, I suggest some theoretical extensions to the work on [research projects](#) in the energy sector. Similarly, the results of the learning model shown in Chapter 6 are discussed in terms of the limitations of the [ESME](#) model and the model formulation outlined in Chapter 5. In Section 7.4.4, I outline extensions to the learning model. In Section 7.4.6, I discuss the similarity between the restructuring of uncertainties using conditional dependencies in an expert elicitation and the recognition of dynamic uncertainty.

7.4.1 Sensitivity Analysis

I have presented the rationale for the choice of method of the [global sensitivity analysis](#) and how this relates to the data limitations and the nature of uncertainty of the [ESOM](#) used. While the results from the [global sensitivity analysis](#) should not be generalised beyond the

ESME model, the methodology itself is generalisable to the other ESOMs.

The effectiveness of the Method of Morris relies to some extent upon the independence of the sampled trajectories⁷. The approach taken for generating the sample used a brute-force optimisation method, developed during this thesis, to identify the most different trajectories from a pool of randomly generated trajectories. However, while this method is guaranteed to provide the most different trajectories, the method does not scale effectively. In other words, the randomly generated pool of samples is finite and its size is limited by the computational effort of comparing the trajectories. The result is that while the input sample is globally optimal, the *global* aspect of the optimisation is limited. An alternative approach, such as the local optimisation provided by Ruano et al., (2012) can be run using a larger initial pool of trajectories. While it does not guarantee a globally optimal input sample, it may achieve a better result from a larger pool with similar running time⁸.

A more serious limitation is related to this fact that a finite number of inputs to the ESME model were incorporated into the sensitivity analysis. A snowball approach was taken, where the initial hypothesis — that technology costs drive the model — influenced the addition of one set of parameters to the model. This analysis was extended to more and more parameters, but ultimately time and computational constraints restrict the degree of model coverage. The grouping of inputs mitigates the cumulative computational burden of adding parameters, but the extensive (and increasing) development time required to add further inputs of the model, which were not initially designed to be permuted computationally, restricted the analysis. Of course, the risk of Type II errors is increased by entirely omitting inputs from the analysis, but the ‘rule of small numbers’ indicates that as increasing numbers of inputs are included, the likelihood of the next input being influential decreases exponentially.

It would be possible to use an iterative approach to improve upon the method taken in this thesis. For example an initial screening analysis conducted with the Method of Morris could be used to identify important input variables. Expert elicitation could then be conducted for a subset of these important variables, And a more detailed Sobol’ analysis used to investigate pairwise interactions.

Another disadvantage of the sensitivity techniques used in this thesis is that correlations between inputs were not taken into ac-

⁷ Experiments with a test function and multiple biased trajectories gave similar results indicating that the approach is not overly sensitive to differences in the number of times parameter levels appear in the input sample. This is probably due to the averaging of the elementary effects.

⁸ The method of Ruano et al., (2012) has now been implemented into the open-source sensitivity analysis library SALib used during this thesis. Unfortunately, this was unavailable during the analysis phase of the work.

count. The Method of Morris enforces the assumption of independence between inputs. However given the purpose for which the sensitivity analysis results were put to, i.e. identifying critical and endogenous uncertainties, this is not a cause for concern. However, correlations are likely to play an important part in some of the exogenous uncertainties such as oil and gas prices, both of which were identified as being critically important to the energy system cost. Although the recommendation in this thesis, that these parameters are treated using scenario analysis would mitigate the effect of this limitation.

7.4.2 *Evaluation of Critical Uncertainties*

The results from the learning model are a function of the revenue values which define the financial benefit of the success of a [research project](#) (or [stage](#) within a [research project](#) for multi-stage projects). In this thesis, the [revenue function](#) is populated by running the [ESME](#), an energy system model of the United Kingdom. The model was run in 10-year time steps, and computed a pathway from 2010 to 2050. The decadal time-steps are likely to have some effect upon the results from the energy system model. A more temporally detailed model would enable a better representation of the pattern of investments made under different input assumptions. Using a greater number of key-years, for example to explore the benefit (if any) of early resolution of an uncertainty would result in a more detailed [revenue function](#).

7.4.3 *The Learning Model*

The aim of the learning model is to quantify the value of critical uncertainties to the energy system. Through the exploration of the strategies available to the decision maker regarding the timing and choice of investment into these [research projects](#) the model indicates the project cost thresholds and probabilities of success within which investment should go ahead. There are number of areas where this model, or approach, could be improved. This can be divided into the following areas:

- computational challenges
- data availability
- model formulation and solution methods

These are now tackled in the following three sections.

7.4.3.1 *Computational challenges*

The results shown in Chapter 6 are limited by the assumption of two key-years in the energy system model, the analysis of two critical un-

certainties that are amenable to research and the division of the projects into two *stages*. The barriers preventing investigation in greater detail are both computational and time constraints. Equation 40 in Chapter 5 shows how the number of scenarios in the energy system model increases as a function of the number of projects and the number of *stages* into which the projects are divided. This means that further insight can be achieved through intensive scenario runs of the energy system model.

7.4.3.2 *Data availability*

The learning model uses the energy system model to generate a *revenue function* which influences the choice and timing of investments into *research projects*. The outcomes of research and development into new energy system technologies, resources and infrastructures, would likely have benefits outside of the energy system. These are not accounted for in the *revenue function*. On the other hand there are a number of perceived barriers to the learning projects which are included in this study. For example, the availability of domestic sustainable biomass, is constrained within the UK due to finite land area and competing uses for land, including agriculture. Even if the biomass availability *research project*, loosely defined as this is in this study, were to succeed, the use of sustainable biomass may still not be able to go ahead due to these barriers.

The *research project* outcomes are based on chosen values within ranges from the literature. For example, in the two-*stage* biomass availability project, the first and second *stages* correspond in a linear increase starting at zero TWh/year and ending at 180 TWh/year. This choice of research outcomes directly influences the research value for the first and second *stages*.

The probabilities assigned to both the individual stages and whole research projects are important in determining the threshold, under which investment in those and other projects (due to interactions between the project outcomes) should proceed. Determining a source for these probabilities is of key importance if this approach is to have further use outside of academia.

Widening the data issues out to the whole *ESME* model, the *revenue function* which defines the value of the *research projects* is a function of the package of assumptions in the *ESME* model. These results are entirely dependent upon the assumptions incorporated into the *ESME* model, including its data and formulation.

7.4.3.3 *Model formulation and solution methods*

The learning model assumes that *research projects* are sequential, and that the outcomes from the projects are contained within the set of [pass, fail]. However, a more realistic representation of *research*

projects could include the partial benefits that could even arise from an otherwise failed project. There is also no way to represent knowledge spillovers, or clustering effects within this framework. More complex dependencies between multiple projects⁹ cannot be incorporated without extending the model formulation, although this would be relatively easy to do within the existing framework.

The approach requires that the starting point of the research programme is the most expensive option, with each stage reducing the cost. However, there may be cases where the most likely estimate is then shown to be incorrect, through the resolution of a research project, with the observed value significantly more expensive than was expected. This would result in an avoided loss, but is difficult to frame within the approach here.

7.4.4 *Extensions to the Learning Model*

The modelling approach encapsulated in the learning model took the formulation applied to sequential research projects in the pharmaceutical sector and adapted it for projects in the energy sector. There is scope to complete the adaptation of the model to make it even more suited to energy research projects. Some of the more theoretical ideas are mentioned in Section 7.4.5, and in this section, I focus on the potential for technical improvements to the model.

7.4.4.1 *A Real Option Approach*

Although the formulation differs between the learning model, (a multi-stage stochastic mixed-integer programme) and a real-options approach, such as that in Childs et al., (1999), the insights that are obtained are similar. For example, the multi-stage stochastic programming formulation of the learning model can enable the option values of decisions to be computed and compared. The value of flexibility explored in Section 5.1.2 is similar in nature to the elaboration of real-options over NPV approaches as outlined in Dixit et al., 1994. The learning model formulation also takes into account the compound nature of the real-options for a portfolio of sequential research projects in isolation or in parallel, computing the optimal investment strategy over all possible outcomes of the projects. However, the connection between the model formulation as a stochastic programme and real-options theory is left for a keen financial mathematician.

The approaches differ in the following ways. The focus of real-options tends to be more financial — investigating the role of volatility in the market upon R&D investments, often in a continuous, rather than discrete settings. The learning model is more straight-forward,

⁹ For example: project A must be completed successfully before project B and C can start. Or: both projects X and Y must complete before commencing Z.

operating over a discrete time horizon, and with the linkage to [ESME](#), a discrete revenue function from the combinations of successful outcomes of [research projects](#) on the energy system.

Another difference is in what is used as an input, and produced as an output. For example, the formulation in Childs et al., (1999) takes a (sub-)set of possible research [strategies](#) as inputs, which are then evaluated to find the optimal under a set of system conditions. In the learning model, the strategies are an output of the model while the costs and value of the successful project are inputs. As the results show, the number of strategies increases exponentially, depending upon the number of [research projects](#) and [stages](#). The exploratory nature of the learning model results may be useful in circumstances where the set of inputs are available.

Fundamentally, the approaches are similar, and which approach is used depends upon the following factors. A real-options approach using dynamic programming allows complex phenomena to be modelled (such as increasing the rate of learning), but requires the use of bespoke algorithms to solve for anything other than the simplest of problems. In the setting of R&D projects, a small number of alternative investment strategies are provided as inputs, with thresholds for cost and probability given as outputs. A learning model approach using multi-stage stochastic mixed integer programming is less flexible in the specification of the model (rate of learning is fixed, research projects have fixed durations), although these simpler mechanisms can be combined to model more complex phenomena (could model mutually exclusive projects through resource constraints). However, commercial mixed-integer solvers can be used to find a globally optimal solution. In the R&D setting, research strategies are given as an output, while the parameters of individual projects are given as inputs.

7.4.4.2 *Endogenous and exogenous uncertainty*

The sensitivity analysis in Chapter 4 revealed that several of the most sensitive inputs to the energy system model [ESME](#) are exogenous uncertainties. Notably the price of natural gas and the price of liquid fuel are two critical parameters for the model results. It would be particularly interesting to introduce exogenous uncertainties into the learning model. This is because several of the endogenous uncertainties, of which two are included in this analysis, interact with the exogenous uncertainties. For example, in the [ESME](#) results for domestic biomass, we can see you that any increase in biomass is correlated with a decrease in the use of liquid fuels. Similarly an increase in the use of domestic biomass reduces imports of biomass the price of which is classified as an exogenous uncertainty. This example shows how the use of domestic biomass potentially decreases exposure to variation in these exogenous uncertainties. On the other hand the use

of CCS means that fossil fuels can continue to be used even in low carbon scenarios. In the results shown in Chapter 6 CCS extends the use of natural gas. And as we know from the [global sensitivity analysis](#) in Chapter 4, the [ESME](#) model is very sensitive to the price of natural gas resources.

Both of these examples provide a strong incentive to investigate the effect of both exogenous and endogenous uncertainties together in the learning model. It may be that biomass seems more attractive once the interactions with exogenous uncertainties taken into account. And the opposite may be the case for CCS which may increase the exposure of a low carbon energy system to fluctuating fossil fuel prices.

This has been left for future work as the integration of both exogenous and endogenous uncertainties would greatly complicate the formulation of the learning model. The work of Dupacova, (2006) outlines some of the issue with this approach at general level. In the specific case of this thesis, the use of the exogenous uncertainties would require many more scenarios to be run in the [ESME](#) model to construct a [revenue function](#).

The mixed-integer stochastic programming formulation used in this work can accommodate up to seven [five-stage](#) projects before special solution techniques required. The current formulation of the model is sufficient for significantly more complex combinations of [research projects](#) than those explored in this thesis.

There is also scope to integrate the learning model with the energy system model. This would allow both endogenous and exogenous uncertainties to be traded off against one another, with research investments essentially used as mitigation options. This is likely to dramatically increase the solution time of the energy system and learning model linkage but could go towards providing so interesting results.

7.4.4.3 *Improvements to the Algorithms for Scenario Generation*

The algorithms used to generate the scenarios in the [ESME](#) model to compute the [revenue function](#) could be made more efficient. For example, Equation 40 and the related combinatorial algorithm do not take into account the duration of a [research project](#). This could lead to extra scenarios being run for combinations of [stages](#) that cannot be completed before the key-year.

Another improvement could be to run the [ESME](#) scenarios as part of a branch-and-bound solution algorithm for the learning model, so that only those scenarios that are required are run.

And finally, there learning model can be run independently of the [ESME](#) model, so could be used on its own, with a modified [revenue function](#), to compute optimal research strategies under uncertainty.

7.4.5 Energy Research Projects

The fundamental unit within the learning model is the [research project](#). The definition of what a [research project](#) constitutes was kept intentionally vague, as the focus of this thesis was on identifying and valuing decision dependent uncertainties. However, there is definitely scope to implement a more refined view of the research process. It could be possible to differentiate between blue-sky research, lab-based research and [commercialisation](#) efforts using the existing model parameters. These parameters include project cost, resource constraints, duration and probability of success. The following sections discuss some more theoretical aspects of [research projects](#).

7.4.5.1 Aspects of technology research-projects

An alternative to the systematic approach developed during the course of this thesis, would be to first identify [research projects](#) of interest, and categorise their pertinent aspects. These aspects could then be used to explore the relative advantages or disadvantages of research into particular types of projects. For example, novel low-carbon technologies could respond to investment with rapid initial, but diminishing cost reductions, but be deployed only in niches. An established technology may have widespread take-up, but only respond to a research project with incremental improvements. In an energy R&D context, this was touched on by Baker et al., (2006, 2008) who modelled at a broad scale incremental versus breakthrough energy technologies. Their results at the global scale were ambiguous, indicating that the outcomes from research were case-specific. Of the technology research aspects outlined below, *granularity* and *novelty* are innate to the technologies themselves and are mostly known before starting a [research project](#). *Structure* and *flexibility* are aspects of the technology [research project](#) that are influenced by the *granularity* and *novelty* of the technology, but can also be modified by the decision maker. *Granularity* places an upper bound on the *flexibility* of a project. The *novelty* of a technology influences the pattern of *payoffs* relative to time and research effort input as discussed in Section 7.4.5.2. While the direct-effect (such as an increase in efficiency or a reduction in cost) of a technology [research project](#) may be well understood, the *payoff* from a project is less certain, due to *interaction* effects with other technology [research projects](#), and also due to the exogenous uncertainties associated with the future development of the energy system.

GRANULARITY Granularity is the resolution to which a technologies' [research project](#) can be divided into separate consecutive [stages](#). The resolution may be naturally constrained by physical, financial or practical aspects of a technology. Granularity also depends upon the scale of what is defined as a '[research project](#)'. For example, a

research project for CCS in the UK, may take as its smallest divisible part an individual prototype CCS power plant, which may still represent several hundred MW and tens of billions of pounds investment. In contrast, a solar photovoltaic research programme has a much fine granularity, potentially at the scale of individual watts for solar cells.

NOVELTY Novelty refers to the maturation of the technology through the innovation process. Technologies closer to deployment may respond to scientific research in a different way to less mature technologies.

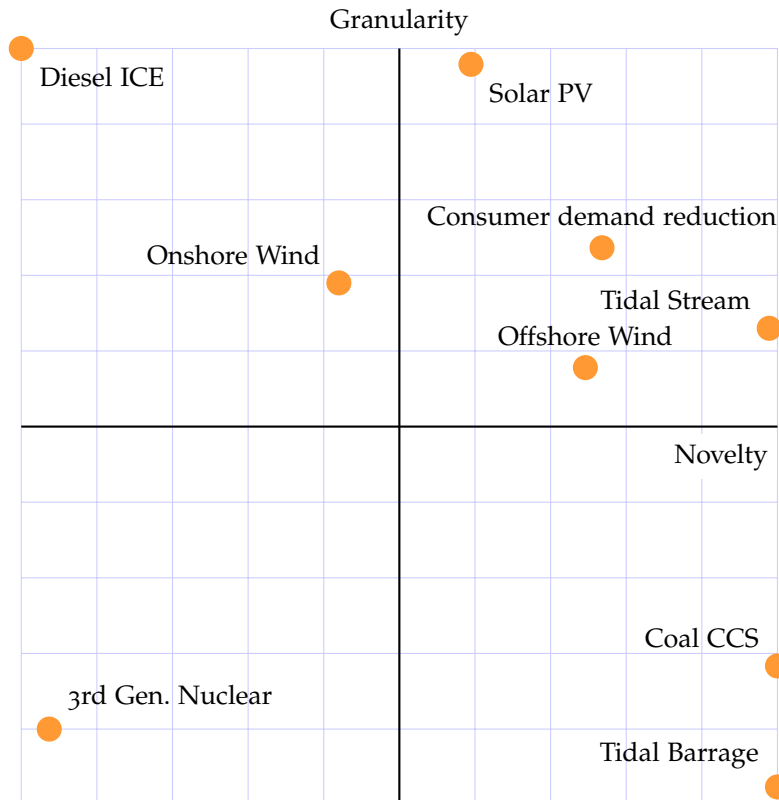


Figure 45: A schematic of energy system components according to novelty and granularity for research purposes

PAYOFF The payoff is the immediate revenue attributable to the successful completion of a phase of a **research project**, or to the successful completion of the final phase of a **research project**.

It is difficult to express the financial payoff of a technology **research project** independently of the system in which that technology will exist. On the other hand, it may be easier to express the **outcome** of a **research project** in terms of the direct-effect of the project — for example an efficiency improvement or a reduction in cost. Anadon et al., (2016) discuss the lack of consensus when eliciting expert judgements of the linkages between the amount invested in research pro-

jects (i.e. the size of the project) and the outcomes of those projects. In addition, the payoff of a project may have financial implications outside of the domain (see spillover, below) which makes it difficult to quantify. Santen et al., (2016) notes that the current availability of data linking the successful outcome of research projects to technology cost reductions is a real constraint on modelling efforts. Baker et al., (2015) discuss the use of energy system models to investigate the payoff from successful research projects. Future work could look at the cost reductions that came out of many expert elicitations to think about how relevant these are to the payoff of successful research projects.

INTERACTION The interactions of the outcomes of different [research projects](#) may not be immediately apparent, but could have important implications for the selection of projects. In the same way that a portfolio approach values diversity of stock returns, a robust strategy for choosing [research projects](#) should take into account the likelihood and independence of returns from [research projects](#).

In the model formulation included in this thesis, the interaction between [research projects](#) incorporated into the cost function which is derived from the output of the [ESOM](#).

FLEXIBILITY The flexibility of a [research project](#) determines how easy it is to abandon the project once the decision to begin the project has begun. If a project is divided into multiple phases, then there are as many opportunities to halt the research as there are phases. If the project has only one phase, then the project continues for the duration, and there is not opportunity to abandon the project early if signs of failure are apparent.

STRUCTURE Structure encompasses the combination of i) the order of phases, ii) likelihood of those phases succeeding, and iii) the payoff associated with each of the phases.

It may be possible to restructure a [research project](#) so that the balance of likelihood of success and payoff is different over the phases. In this case, it is useful to understand when a particular structure is preferable (or even optimal), why it is preferable (optimal), and if there are any general rules or guidelines that can make it easier to structure a [research project](#).

In cases where multiple projects are considered, the structure of a research programme would also encompass the interactions between successful project outcomes, the timing of individual phases of multiple projects as well as the order of phases, likelihood of success and payoffs.

7.4.5.2 *Research archetypes*

In this section, I outline five separate archetypes for [research projects](#). The aim of the archetypes is to provide a descriptive categorisation to [research projects](#) and provide a representation of desirable attributes of the balance of benefits and costs of a [research project](#) over time. This list of archetypes is not claimed as definitive or declarations of some kind of physical law, but are intended to foster easier discussion of [research projects](#).

- moon shot
- big pharma
- low hanging fruit
- linear learning
- accelerated learning

and two confounding effects:

- negative learning
- spillovers

MOON SHOT An all-in project, in which there is only one phase, the successful completion of which results in massive revenue. However, the likelihood of success is low. The work of Dixit et al., (1994) suggests that a project structured in such a crude way is less valuable than one that is phased, as the value of flexibility (the option to continue or not) is ignored.

BIG PHARMA A project is split into consecutive phases, but payoff only occurs on successful completion of final phase. As described in Colvin et al., (2008), the pharmaceutical R&D pipeline consists of five [stages](#) ranging through

- discovery
- preclinical trials
- clinical trials - three phases

with successful completion of these three phases resulting in regulatory approval and [commercialisation](#).

LOW-HANGING FRUIT Project is split into consecutive phases. The initial [stages](#) result in large revenues for relatively small investment, with decreasing returns to investment. This is the same profile as a learning curve, where incremental but diminishing returns are realised in proportion to cumulative R&D budget.

LINEAR LEARNING Project is split into consecutive phases. The relationship between investment and revenue is linear. This is a special case of a learning curve, where the gradient of the learning is constant.

We can also describe **research projects** by the balance of expected net revenue across the project. If revenue is divided equally among three phases of decreasing probability of success, and given that the probabilities accrue according to Equation 38, the probability of success of each phases is conditional upon the successful completion of the previous phase, this probability of success for the final phase can be very small indeed. If probabilities increase across phases, the smaller probability of the first research phase also has the effect of diminishing the expected net revenue of the final phases, discouraging investment in the entire project. Under decreasing revenues across phases, a corresponding increase in revenue is required to offset the rapid reduction in probability of success.

CONFOUNDING EFFECTS: SPILLOVER The R&D project fails its intended task of reducing cost or improving technical performance. However, the project does result in other unintended benefits (spillover benefits) which have tangible value. Spillover provides an overall incentive for investment in R&D.

CONFOUNDING EFFECTS: NEGATIVE LEARNING Negative learning has several parallel definitions in the literature. Oppenheimer et al., (2008) states that it is the process whereby new technical discovery leads to beliefs that move away from the *a posteriori* correct answer. In the context of this work, negative learning implies that R&D results in a belief that is more wrong than before the R&D was conducted. For example, the successful result of an R&D project may suggest that the cost of a technology is £100/MWh, but upon construction the cost is found to be £150/MWh. This is equivalent to overestimating the cost reduction associated with the successful **outcome** of a project, making R&D investment decisions based upon this overconfident belief, and then only discovering of this overconfidence upon deployment of the technology. However, the treatment of negative learning works better with the Type I endogenous uncertainty (see Table 12) which is not investigated in this thesis.

7.4.5.3 *Assigning research archetypes to components of the energy system*

The outcomes of **research projects** can be directed not only energy technologies, but also at other components of the energy system. Enabling infrastructures, such as power lines and pipelines, so called 'smart' technologies, such as automated and 'clever' heating controls, metering and demand management technologies, could also have a tangible economic and social benefit to the future energy system.

7.4.6 Eliciting Uncertain Input Data for the Research Projects

Eliciting the uncertainties in this study would involve a linear, but protracted process. The first step is identifying influential parameters. Then, evaluating the costs associated with resolving an uncertainty. And finally transforming the uncertain parameters from a plausible range of unknown likelihood to a project of multiple uncertain phases associated with a known parameter value. For example, an influential uncertainty could be “The availability of CCS in 2030.” Without recognising the nature of the uncertainty as decision dependent, a probability distribution associated with such a vague uncertainty is very difficult to quantify. However, concentrating on a series of consecutive research [stages](#), each of which is associated with a probability of success, a cost and duration, would ease this, by expressly recognising the influence of the decision to begin each [stage](#). While a number of elicitation have focussed on technology cost, an expert elicitation by Anadon et al., (2012) asked expert about the cost and timing of Generation IV nuclear reactor availability.

This is a form of *structuring*, as it is known in expert elicitation (O’Hagan, 2006a). Structuring involves moving from an uncertainty which is perhaps too complex for elicitation of uncertainties, to one that is more easily unpacked and understandable. This is normally done through division of an uncertainty into a package of conditionally dependent uncertainties.

For example, the elicitation of a probability distribution for the CCS example above, requires the expert to integrate their belief for whether the policies and actions will be made that will result in CCS, as well as the technology itself working. In other words, this uncertainty is expressed as exogenous. By conditioning the uncertainty on the decision to invest in a particular project, the uncertainty becomes “What is the likelihood of a £10bn project [stage](#) successfully resulting in the availability of CCS in 2030?” This is equivalent to the decision dependent uncertainty, and one that is conditioned upon action. This is easier to quantify than the exogenous example above.

The following sections cover the established method and formulation for conducting a screening analysis using the Elementary Effects Method, also known as the Method of Morris.

While the work reproduced here does form an important part of the methodology of the thesis, the formulation for elementary effects is derived from existing work, primarily that of Campolongo et al., (2007), Morris, (1991) and Saltelli et al., (2008a). However, the thesis does provide a small extension to previous work by offering the formulation of the combinatorial optimisation problem first suggested in Campolongo et al., (2007). As original work, this extension is outlined in Section 3.3.1.

This appendix chapter consists of four sections, which cover in turn the formulation for generating samples, the grouping of input parameters, computation of elementary effects and finally computing the sensitivity indices.

A.1 SAMPLING METHOD

Given a model with k independent inputs $X_i, i = 1, \dots, k$, each of which is divided into p discrete **levels** to produce a grid Ω , the objective of the sampling method is to generate a panel of N trajectories. N is an exogenously defined sample value and defines the total size of the sampling procedure and interacts with the number of **levels** p into which the inputs are divided. p defines the ‘resolution’ of the analysis, with higher values of p (higher resolution) requiring a bigger input sample for the same degree of confidence. As suggested in Saltelli et al., (2008a), a p of 4 requires an N of 10 or more.

Each trajectory is represented here by matrix \mathbf{B}^* and has dimensions $(k + 1) \times k$ which provides one randomly selected elementary effect per input parameter:

$$\mathbf{B}^* = (\mathbf{J}_{k+1,1} \mathbf{x}^* + (\Delta/2) [(2\mathbf{B} - \mathbf{J}_{k+1,k}) \mathbf{D}^* + \mathbf{J}_{k+1,k}]) \mathbf{P}^* \quad (49)$$

where \mathbf{D}^* is a k -dimensional diagonal matrix in which each element is either $+1$ or -1 with equal probability, \mathbf{P}^* is a k -by- k random permutation matrix in which each row contains one element equal to one, all others are zero, and no two columns have 1’s in the same position. \mathbf{x}^* is a randomly chosen base value of \mathbf{X} from the p -**level** grid Ω and is thus a vector of dimensions k . \mathbf{x}^* provides an initial starting point for each trajectory, upon which the randomly generated permutation matrix \mathbf{P}^* operates. \mathbf{B} is a $(k + 1) \times k$ matrix with elements that are

either 0 or 1 and has the key property that for every column index $j, j = 1, \dots, k$ there are two rows of \mathbf{B} that differ only in the j th entry. This is fulfilled through using a strictly lower triangular matrix of 1s. $\mathbf{J}_{k+1,k}$ is a $(k+1) \times k$ matrix of 1s. Finally, Δ is computed as a function of the number of levels $p/2(p-1)$ when p is even and denotes the order in which the inputs move in each row of the \mathbf{B}^* matrix.

A.2 WORKING WITH GROUPS

When working with groups of inputs, the sampling procedure needs to be modified with a group matrix \mathbf{G} which identifies membership of the k parameters to \bar{G} groups. $\mathbf{G}(i, g)$ is equal to 1 if parameter i is a member of group g ; otherwise $\mathbf{G}(i, g) = 0$. As all factors in a group move together, the matrix of trajectories \mathbf{B}^* is now of dimension $(\bar{G} + 1) \times k$, and subsequently matrix \mathbf{B} is of size $(\bar{G} + 1) \times \bar{G}$.

$$\mathbf{B}^* = (\mathbf{J}_{\bar{G}+1,1}\mathbf{x}^* + (\Delta/2) [(2\mathbf{B}(\mathbf{G}\mathbf{P}^*)^T - \mathbf{J}_{\bar{G}+1,k})\mathbf{D}^* + \mathbf{J}_{\bar{G}+1,k}]) \quad (50)$$

A.3 COMPUTING ELEMENTARY EFFECTS

The elementary effect of the i th input parameter computed along trajectory j is given by:

$$\text{EE}_i^j(\mathbf{x}^{(l)}) = \frac{[y(\mathbf{x}^{(l+1)}) - y(\mathbf{x}^{(l)})]}{\Delta}, \quad (51)$$

if the i th component of $\mathbf{x}^{(l)}$ is increased by Δ , and

$$\text{EE}_i^j(\mathbf{x}^{(l+1)}) = \frac{[y(\mathbf{x}^{(l)}) - y(\mathbf{x}^{(l+1)})]}{\Delta}, \quad (52)$$

if the i th component of $\mathbf{x}^{(l)}$ is decreased by Δ .

A.4 COMPUTING METRICS

Once the elementary effects have been computed, it is a simple matter to compute the desired sensitivity metrics:

$$\mu_i = \frac{1}{r} \sum_{j=1}^N \text{EE}_i^j \quad (53)$$

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^N |\text{EE}_i^j| \quad (54)$$

$$\sigma_i^2 = \frac{1}{r-1} \sum_{j=1}^N (\text{EE}_i^j - \mu_i)^2 \quad (55)$$

Note that for the treatment of groups, as used in this paper, the computation of μ^* is altered slightly so that it is indexed by group g , rather than by parameter i :

$$\mu_g^* = \sum_{i=1}^k G_{ig} \times \mu_i^* \quad (56)$$

and the metrics μ_i and σ_i^2 are no longer applicable.

Instead a confidence interval can be computed, using bootstrapping. First, the elementary effects are re-sampled R times to produce an array of dimensions N -by- R called $\hat{E}\hat{E}_i^r$.

$$\hat{\mu}_i^* = \frac{1}{r} \sum_{j=1}^N | \hat{E}\hat{E}_i^r | \quad (57)$$

$$\hat{\sigma}_i^2 = \frac{1}{r-1} \sum_{j=1}^N \left(\hat{\mu}_i^* - \frac{1}{I} \sum_{i=1}^I \hat{\mu}_i^* \right)^2 \quad (58)$$

The confidence interval CI is then computed from the mean and standard deviation of the average of the absolute re-sampled elementary effects, where C is the inverse of the confidence bound e.g. 5% for a 95%CI.

$$CI = \hat{\mu}^* \pm \frac{(1-C)}{2} \times \frac{\hat{\sigma}_i^2}{\sqrt{R}} \quad (59)$$

THE STOCHASTIC MIXED INTEGER PROGRAMME

The formulation of Colvin et al., (2008) is replicated below, with some modification to the notation to match the remainder of this thesis.

The model is formulated as a mixed-integer multi-stage stochastic programme. In §B.1 we describe the decision variables for the R&D model. In §B.2 we describe the non-anticipativity constraints, and in §B.3 we show the remainder of the formulation.

B.1 DECISION VARIABLES

The decision variable Y_{ijts} is defined entirely in terms of variable X_{ijts} using the formulation from Colvin et al., (2008):

$$Y_{ijts} = \sum_{t'} X_{ijt's} \quad \forall i, j, t' < t + 1 - \tau_{ij, s} \quad (60)$$

The variable Q_{ijts} is used to compute revenues for those R&D projects not completed at the end of the time horizon:

$$Q_{i,I,1,s} = 1 - X_{i,I,1,s} \quad \forall i, s \quad (61)$$

$$Q_{i,I,t,s} = Q_{i,I,t-1,s} - X_{i,I,t,s} \quad \forall i, t > 1, s \quad (62)$$

$$Q_{i,j,t,s} = Q_{i,j,t-1,s} + X_{i,j-1,t-\tau_{i,(j-1)},s} - X_{i,j,t,s} \quad \forall i, j > I, t > 1, s \quad (63)$$

The objective function is to maximise the total expected net revenue from the R&D pipeline:

$$\text{Obj} = \sum_s p_s (\text{Rev}_s^T - \text{Cost}_s^T) \quad (64)$$

The financial variables, total revenue Rev_s^T and total cost Cost_s^T are computed as follows:

$$\text{Rev}_s^T = \begin{cases} \sum_{i,t} (d\text{Rev}_{i,t+\tau_{ij}} \cdot X_{i,j,t,s}) & \forall (i, s) \in I_{is}^S, t \\ \sum_{i,j,t} (d\text{Rev}_{i,t+\tau_{ij}} \cdot X_{i,j,t,s}) & \forall (s, i, j) \in R_{sij}, t \end{cases} \quad (65)$$

$$\text{Cost}_s^T = \sum_{i,j,t} \left((1-r)^{(t-2010)} \cdot \text{Cost}_{ij}^R \cdot X_{i,j,t,s} \right) \quad \forall i, j, t \quad (66)$$

B.2 NON-ANTICIPATIVITY CONSTRAINTS

The non-anticipativity constraints force decision variables $X_{i,j,t,s}$ to the same value over s , only if scenarios differ by at most one outcome, as defined by the subset $\text{sdif}_{s,s'}$. The **research projects** and **research stages** which differ between this subset of scenario pairs are defined by the the subsets $i^{s,s'}$ and $j^{s,s'}$.

$$\text{sdif}_{s,s'} = \begin{cases} 1 & \forall (s, s') \in \sum_i^I (\text{oc}_{s,i} - \text{oc}_{s',i}) = 1, s > s' \\ 1 & \forall s' > s \in \sum_i^I (\text{oc}_{s',i} - \text{oc}_{s,i}) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (67)$$

Non-anticipativity constraints are active until the completion of the differentiating **stage** ($i^{s,s'}, j^{s,s'}$). The constraints can be expressed in terms of the decision variable $Y_{i^{s,s'}, j^{s,s'}, t, s}$.

$$-Y_{i^{s,s'}, j^{s,s'}, t, s} \leq X_{i,j,t,s} - X_{i,j,t,s'} \leq -Y_{i^{s,s'}, j^{s,s'}, t, s} \quad \forall i, j, (s, s') \in \Psi, t > 1 \quad (68)$$

B.3 OTHER CONSTRAINTS

The first project **stages** should be identical.

$$X_{i,1,1,s} = X_{i,1,1,s'} \quad \forall i, s \quad (69)$$

Research project stages should only be performed once.

$$\sum_t^T X_{i,1,1,s} \leq 1 \quad \forall i, j, s \quad (70)$$

Research project stages must be performed consecutively.

$$\sum_1^{t' \leq t} X_{ij,t's} \leq Y_{i,j-1,t,s} \quad \forall i, j > 1, t > 1 \quad (71)$$

Forces abandonment of a **research project** if a previous **stage** fails.

$$X_{ijts} = 0 \quad \forall s, i, j \in f, t \quad (72)$$

Research project stages should not begin until the completion of the previous **stage**.

$$X_{ijts} = 0 \quad \forall (s, i, j) \notin f, t < (1 + \tau_{i,j}) \quad (73)$$

A resource constraint is imposed to ensure that demand for resources to conduct [research projects](#) does not exceed supply.

$$\sum_i \sum_j \sum_{t' \geq (t - \tau_{i,j} + 1)}^{t' \leq t} \rho_{i,j,k} \cdot X_{ijt's} \leq \rho_k^{\max} \quad \forall k, t, s \quad (74)$$

UNCERTAINTY FRAMEWORKS: RISK, UNCERTAINTY, AMBIGUITY, IGNORANCE

There is a very large literature on uncertainty in policy-relevant, model supported decision making under risk and uncertainty. Stirling, (2010) illustrates the importance to scientific integrity of retaining the plurality and conditional nature of knowledge in communication with policy makers. A narrow focus on *Knightsian Risk*, in which a reductionist approach, often through application of quantitative models, produces a definitive interpretation of a problem is potentially damaging to the scientific-policy process as well as misleading. This is because a blinkered *Knightsian Risk* approach displaces the recognition of other more insidious types of uncertainty which require different approaches. These approaches are depicted in Stirling's useful matrix (Stirling, 2007), which relates the type of uncertainty to the degree of 'certitude' in both likelihoods and effects. It is reproduced here in Figure 46. The earlier typology of Wynne, (1992), reproduced in Spiegelhalter et al., (2011a), is similar to Stirling's matrix. Wynne, (1992) differentiates between risk, uncertainty, ignorance and indeterminacy. Stirling's matrix expands on Wynne, (1992) by explicitly introducing the two dimensions of possibility and probability although neglects Wynne's note that uncertainties can exist within more than one of his categories at the same time. Wynne, (1992) also points out the explicitly conditional nature of science, in that to structure a rigorous scientific study, it is always necessary to place some uncertainties outside a study. Spiegelhalter et al., (2011a) discuss uncertainty in formal models for decision making under risk, rather than Stirling's focus on scientific uncertainty in policy making. They offer a five-level approach, identifying the location and source (rather than objects) of uncertainty.

1. future events
2. model parameters
3. model structure
4. acknowledged inadequacies
5. unknown inadequacies

As such, they expand upon Wynne, (1992) to include the questions surrounding the capacity of a model to answer a given question using a similar onion style diagram to Rotmans et al., (2001a). Thus,

any uncertainties in a higher level (such as model-question suitability) should take precedence over a lower level (such as forecast uncertainty). Keynes' interpretation of this same problem is outlined in Dow, (2004).

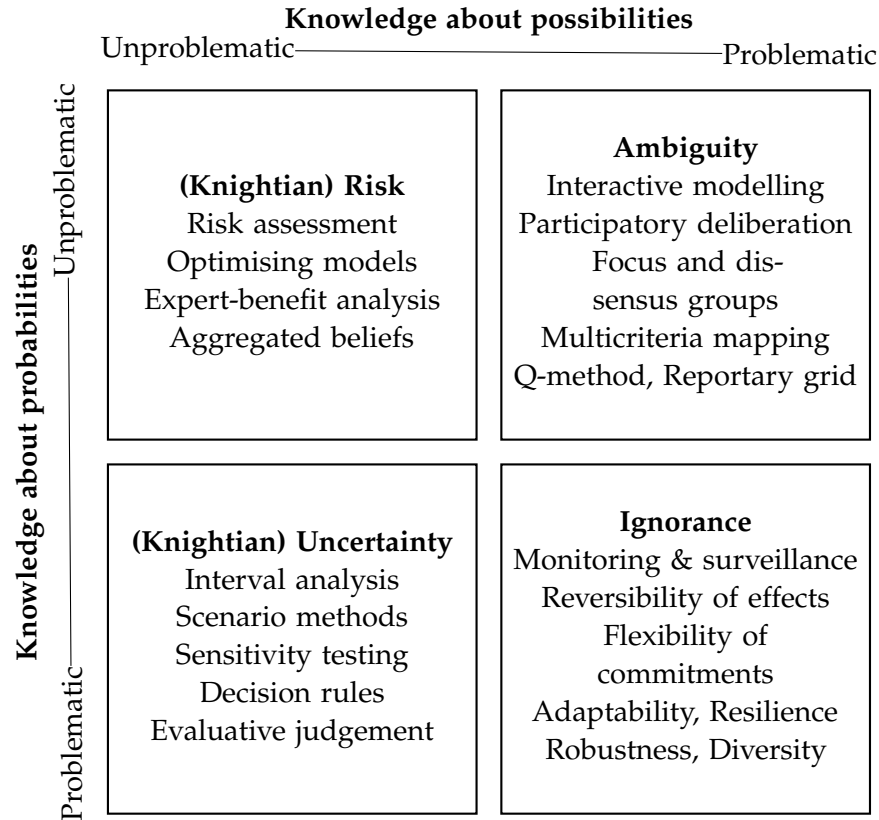


Figure 46: Methodological responses to different forms of incertitude (adapted from Stirling, 2007, 2010)

McManus et al., (2006) offer an engineering and systems design perspective on uncertainty. They offer a holistic framework whereby uncertainties and derived risks (note that this is a definition of **risk** distinct from **Knightian Risk**) can also be viewed as opportunities. Risks and opportunities can be mitigated or exploited in turn. The resulting outcome has desirable characteristics including **robustness**, **adaptability**, **evolvability** and **flexibility**. These outcomes match Stirling's methodological responses to ignorance.

A distinct effort to produce a taxonomy of uncertainty has been undertaken in the Netherlands **IAM** community (Asselt, 2000; Asselt et al., 2002; Rotmans et al., 2001a,b; Sluijs, 1997; Walker et al., 2003). The typology in Rotmans et al., (2001a,b) draws upon that in Asselt, (2000). They identify sources and types of uncertainty in **IAMs** and integrate these into a framework. Their uncertainty types, later adopted, extended and superseded by Spiegelhalter et al., (2011a), are a useful contribution to the literature. Under conditions of **Knightian**

Uncertainties	Risk / Opportunities	Mitigation / Exploitation	Outcomes
Epistemic Inexactness Variability Aleatory Ambiguity Ignorance	Disaster Failure Degradation Schedule (±) Market shifts (±) Need shifts (±) Extra capacity Emergent capabilities	Margins Redundancy Design Choices Verification and Test Generality Upgradeability Modularity Tradespace Exploration Portfolios & Real Options	Reliability Robustness Versatility Flexibility Evolvability Interoperability

Figure 47: Framework for handling uncertainties and their effects (adapted from McManus et al., 2006)

uncertainty Asselt et al., (2002) apply a pluralist approach to uncertainty management in IAMs.

From these papers, key concepts emerge. Conditional upon the task and perspective at hand, uncertainty may be viewed in different ways. Fundamentally, there are two different types of uncertainty (or ‘natures’ to use the terminology of Walker et al., (2003)). The first is epistemic uncertainty or that which occurs due to lack of knowledge. This is likened to not knowing the value of a coin after it has been flipped, but not yet revealed (Spiegelhalter et al., 2011a). Epistemic uncertainty can be reduced through waiting and learning, more detailed modelling of the system under enquiry, or expending effort to learn. The second type of uncertainty is aleatory uncertainty. This is uncertainty due to natural randomness, although there is some debate as to whether this exists in reality. The “... age old argument of determinism. . .” is that no phenomena is truly random, that all uncertainty is epistemic, and that if a detailed enough model were to exist, we could reduce this uncertainty (O’Hagan et al., 2004). However, this argument is subject to infinite regress. Perhaps aleatory uncertainty is a useful concept from a modelling perspective, because it is easier to view a process that operates well below the resolution at which one is modelling as a truly random process (using parameterisation for example). Finally, the key distinction between epistemic and aleatory uncertainty is that the former can be reduced, while the latter is irreducible.

Comparing Figures 46 and 47 demonstrates the range of interpretations of important concepts such as risk and uncertainty. McManus et al., (2006) use risk and uncertainty in a more general sense than Stirling, (2007, 2010) and Spiegelhalter et al., (2011a), but the links they propose, especially making explicit the relationship between risk and opportunities and mitigation and exploitations, are most useful.

Thus, the frameworks of Spiegelhalter et al., (2011a) and Walker et al., (2003) provide a segue from that of Stirling, (2007, 2010) to the world of quantitative modelling, with McManus et al., (2006) providing a useful counterpoint. Quantifying uncertainty so that we are able to make decisions under risk is very useful for policy makers, but only accurate if we are dealing with uncertainties of a particular nature. Beyond [Knightian Risk](#), we must caveat our results or use deliberately imprecise expressions of uncertainty in our results due to ambiguity, ignorance or Knightian uncertainty. Alternatively, or in addition, we can use a plurality of approaches to provide a number of different perspectives on the same problem, making explicit the need for human judgement to take precedence over modelled insight.

INPUT DATA

D.1 GROUPING OF INPUT PARAMETERS FOR SENSITIVITY ANALYSIS

Table 31: Storage Costs

Group	Technology	Parameter
Battery Capex	Battery - Li-ion	Storage Capital Cost
Battery Capex	Battery - Li-ion	Storage Capital Cost Power
Battery Capex	Battery - NaS	Storage Capital Cost
Battery Capex	Battery - NaS	Storage Capital Cost Power
Battery Capex	Flow battery - Redox	Storage Capital Cost
Battery Capex	Flow battery - Redox	Storage Capital Cost Power
Battery Capex	Flow battery - Zn-Br	Storage Capital Cost
Battery Capex	Flow battery - Zn-Br	Storage Capital Cost Power

Table 32: Resource inputs

Group	Resource	Parameter
Coal Resource Cost	Coal	Resource Cost
Liquid Fuel Resource Cost	Aviation Fuel	Resource Cost
Liquid Fuel Resource Cost	Liquid Fuel	Resource Cost
Nuclear Fuel Cost	Nuclear	Resource Cost
Biomass Resource Cost	Biofuel Imports	Resource Cost
Biomass Resource Cost	Biomass	Resource Cost
Biomass Resource Cost	Biomass Imports	Resource Cost
Gas Resource Cost	Gas	Resource Cost
Geo Heat Resource Cost	Geothermal Heat	Resource Cost

Table 33: Constraints

Group	Technology / Resource	Parameter
Electricity Sector Build Rate	Nuclear	Max Build Rate
Electricity Sector Build Rate	Offshore Wind	Max Build Rate
Electricity Sector Build Rate	Onshore Wind	Max Build Rate
Electricity Sector Build Rate	IGCC Biomass with CCS	Max Build Rate
Electricity Sector Build Rate	PC Coal with CCS	Max Build Rate
Electricity Sector Build Rate	CCGT with CCS	Max Build Rate
Electricity Sector Build Rate	IGCC Coal with CCS	Max Build Rate
CCS Industry Build Rate	IndustryCCS	Max Build Rate
Biomass Max Resource Qty	Biomass	Max Resource Quantity
CCS Build Rate	CCSStations	Max Build Rate

Table 34: Heating Technology Costs

Group	Technology	Parameter
DH Capex	DH for Dwelling (HD, ThE)	Capital Cost
DH Capex	DH for Dwelling (HD, ThG)	Capital Cost
DH Capex	DH for Dwelling (HD, ThM with Retrofix)	Capital Cost
DH Capex	DH for Dwelling (HD, ThM with Retroplus)	Capital Cost
DH Capex	DH for Dwelling (HD, ThM)	Capital Cost
DH Capex	DH for Dwelling (HD, ThP with Retrofix)	Capital Cost
DH Capex	DH for Dwelling (HD, ThP with Retroplus)	Capital Cost
DH Capex	DH for Dwelling (HD, ThP)	Capital Cost
DH Capex	DH for Dwelling (LD, ThE)	Capital Cost
DH Capex	DH for Dwelling (LD, ThG)	Capital Cost
DH Capex	DH for Dwelling (LD, ThM with Retrofix)	Capital Cost
DH Capex	DH for Dwelling (LD, ThM with Retroplus)	Capital Cost
DH Capex	DH for Dwelling (LD, ThM)	Capital Cost
DH Capex	DH for Dwelling (LD, ThP with Retrofix)	Capital Cost
DH Capex	DH for Dwelling (LD, ThP with Retroplus)	Capital Cost
DH Capex	DH for Dwelling (LD, ThP)	Capital Cost
DH Capex	DH for Dwelling (MD, ThE)	Capital Cost
DH Capex	DH for Dwelling (MD, ThG)	Capital Cost
DH Capex	DH for Dwelling (MD, ThM with Retrofix)	Capital Cost
DH Capex	DH for Dwelling (MD, ThM with Retroplus)	Capital Cost
DH Capex	DH for Dwelling (MD, ThM)	Capital Cost
DH Capex	DH for Dwelling (MD, ThP with Retrofix)	Capital Cost
DH Capex	DH for Dwelling (MD, ThP with Retroplus)	Capital Cost
DH Capex	DH for Dwelling (MD, ThP)	Capital Cost
Heat Pump Capex	Heat Pump (Air Source, Hot Water)	Capital Cost
Heat Pump Capex	Heat Pump (Air Source, Space Heat)	Capital Cost
Heat Pump Capex	Heat Pump (Ground Source, Hot Water)	Capital Cost
Heat Pump Capex	Heat Pump (Ground Source, Space Heat)	Capital Cost

Table 35: Technology Costs and Efficiencies

Group	Technology	Parameter
Backstop Capex	Air Capture of CO ₂	Capital Cost
CCGT Capex	CCGT	Capital Cost
CCS Capex	CCGT CCS Retrofit	Retrofit Capital Cost
CCS Capex	CCGT with CCS	Capital Cost
CCS Capex	IGCC Biomass with CCS	Capital Cost
CCS Capex	IGCC Coal CCS Retrofit	Retrofit Capital Cost
CCS Capex	IGCC Coal with CCS	Capital Cost
CCS Capex	PC Coal CCS Retrofit	Retrofit Capital Cost
CCS Capex	PC Coal with CCS	Capital Cost
CCS Capex	SNG Plant (Biomass Gasification with CCS)	Capital Cost
CCS Efficiency	CCGT with CCS	Efficiency
CCS Efficiency	IGCC Biomass with CCS	Efficiency
Demand Capex	Lighting (CFL)	Capital Cost
Demand Capex	Lighting (LED)	Capital Cost
H ₂ Capex	H ₂ Plant (Biomass Gasification with CCS)	Capital Cost
H ₂ Capex	H ₂ Plant (Coal Gasification with CCS)	Capital Cost
H ₂ Capex	H ₂ Plant (Electrolysis)	Capital Cost
H ₂ Capex	H ₂ Plant (SMR with CCS)	Capital Cost
H ₂ Capex	H ₂ Turbine	Capital Cost
IGCC Coal Capex	IGCC Coal	Capital Cost
Nuclear Capex	Nuclear	Capital Cost
PC Coal Capex	PC Coal	Capital Cost
Renewables Capex	Geothermal Plant (EGS) Electricity & Heat	Capital Cost
Renewables Capex	Geothermal Plant (HSA) Electricity & Heat	Capital Cost
Renewables Capex	Geothermal Plant (HSA) Heat Only	Capital Cost
Renewables Capex	Hydro Power	Capital Cost
Renewables Capex	Incineration of Waste	Capital Cost
Renewables Capex	Large Scale Ground Mounted Solar PV	Capital Cost
Renewables Capex	Micro Solar PV	Capital Cost
Renewables Capex	Micro Solar Thermal (non south facing)	Capital Cost
Renewables Capex	Micro Solar Thermal (south facing)	Capital Cost
Renewables Capex	Offshore Wind	Capital Cost
Renewables Capex	Onshore Wind	Capital Cost
Renewables Capex	Tidal Range	Capital Cost
Renewables Capex	Tidal Stream	Capital Cost
Renewables Capex	Wave Power	Capital Cost

Table 36: Transport inputs

Group	Technology	Parameter
Bus H2 Capital Cost	Bus (Hydrogen FCV)	Capital Cost
Car Battery Capex	Car Battery (A/B Segment)	Capital Cost
Car Battery Capex	Car Battery (C/D Segment)	Capital Cost
Car CNG Capex	Car CNG (A/B Segment)	Capital Cost
Car CNG Capex	Car CNG (C/D Segment)	Capital Cost
Car H2 Capex	Car Hydrogen FCV (A/B Segment)	Capital Cost
Car H2 Capex	Car Hydrogen FCV (C/D Segment)	Capital Cost
Car H2 Capex	Car Hydrogen ICE (A/B Segment)	Capital Cost
Car H2 Capex	Car Hydrogen ICE (C/D Segment)	Capital Cost
Car Hybrid Capex	Car Hybrid (A/B Segment)	Capital Cost
Car Hybrid Capex	Car Hybrid (C/D Segment)	Capital Cost
Car ICE Capex	Car ICE (A/B Segment)	Capital Cost
Car ICE Capex	Car ICE (C/D Segment)	Capital Cost
Car PHEV Capex	Car PHEV (Long Range A/B Seg)	Capital Cost
Car PHEV Capex	Car PHEV (Long Range C/D Seg)	Capital Cost
Car PHEV Capex	Car PHEV (Med Range A/B Seg)	Capital Cost
Car PHEV Capex	Car PHEV (Med Range C/D Seg)	Capital Cost
Car PHEV Capex	Car PHEV (Short Range A/B Seg)	Capital Cost
Car PHEV Capex	Car PHEV (Short Range C/D Seg)	Capital Cost
LGV Capex	LGV (BEV)	Capital Cost
LGV Capex	LGV (Dual Fuel Direct)	Capital Cost
LGV Capex	LGV (Dual Fuel Port)	Capital Cost
LGV Capex	LGV (Gas SI)	Capital Cost
LGV Capex	LGV (Hybrid)	Capital Cost
LGV Capex	LGV (Hydrogen FCV)	Capital Cost
LGV Capex	LGV (Hydrogen ICE)	Capital Cost
LGV Capex	LGV (ICE)	Capital Cost
LGV Capex	LGV (PHEV)	Capital Cost

ADDITIONAL RESULTS: OPTIMAL TRAJECTORIES

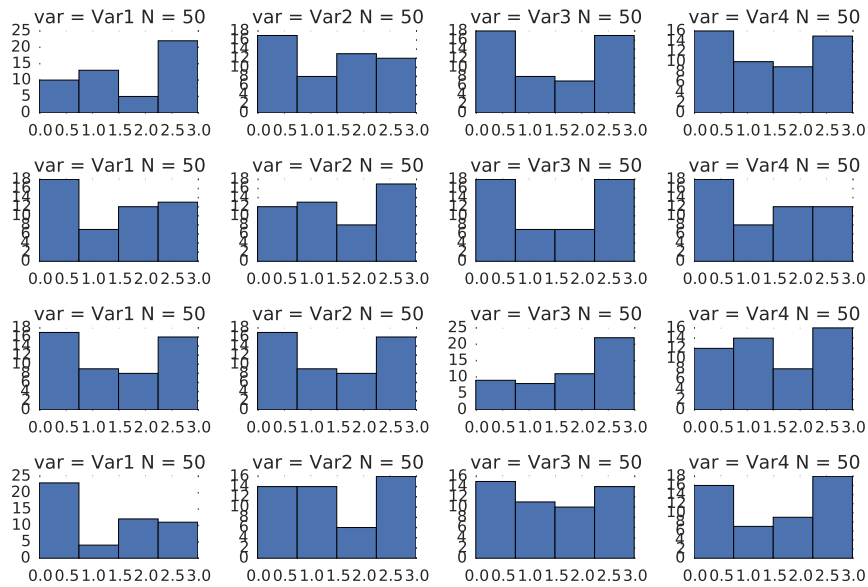


Figure 48: From a pool of trajectories, examples of the best case samples

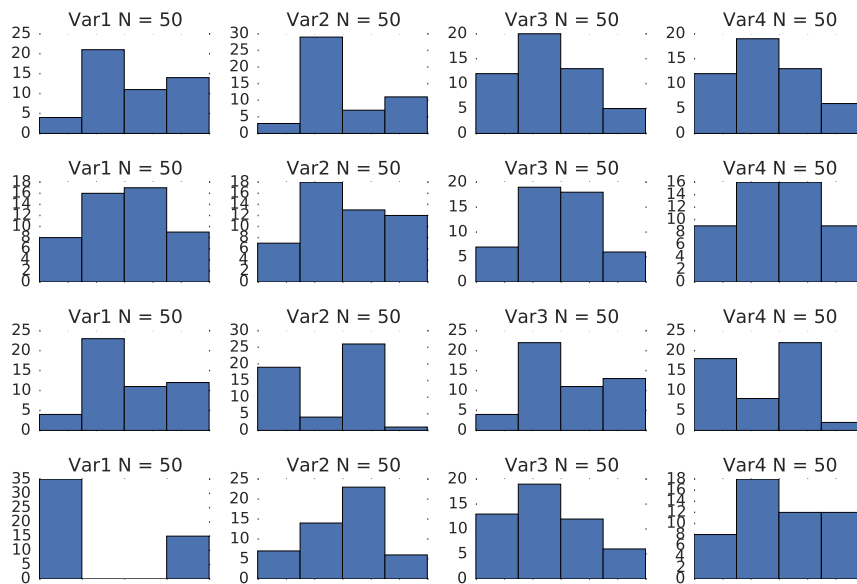


Figure 49: From the same pool of trajectories as above, examples of the worst case samples

ADDITIONAL RESULTS: SENSITIVITY ANALYSIS

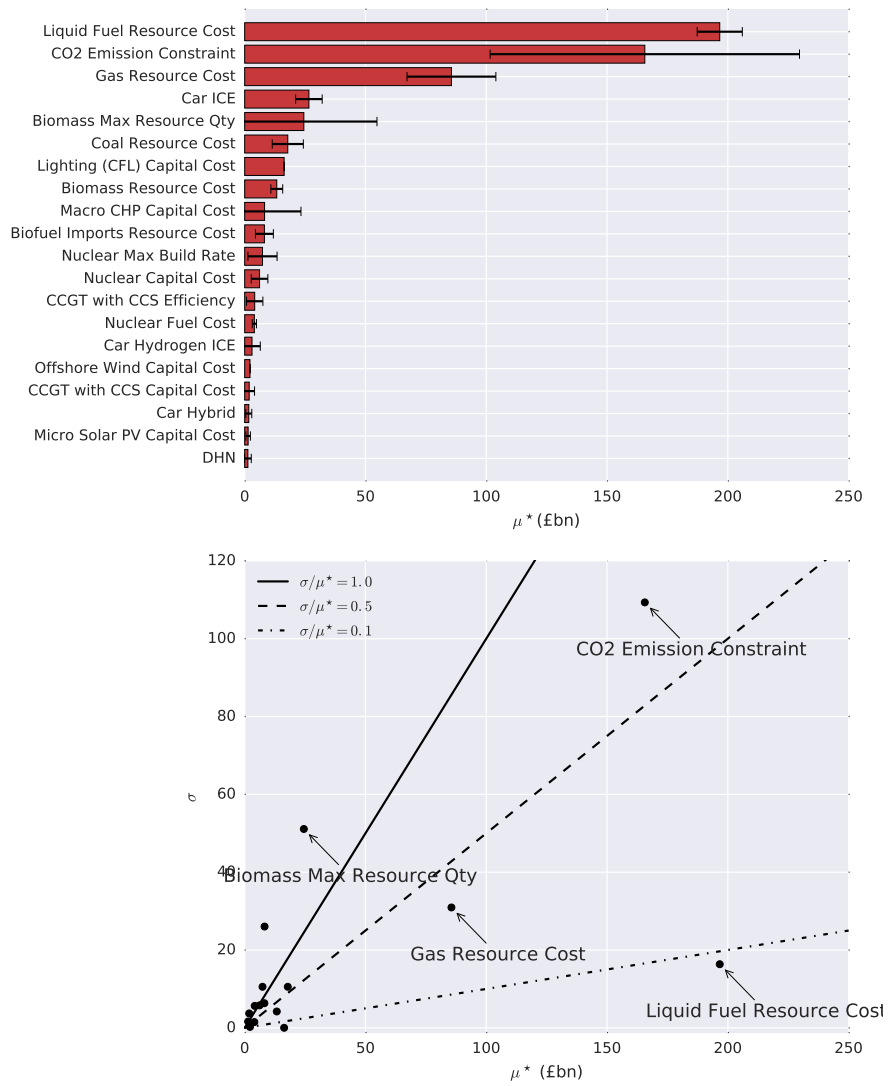


Figure 50: An initial screening analysis indicated that some parameter groups were influential .

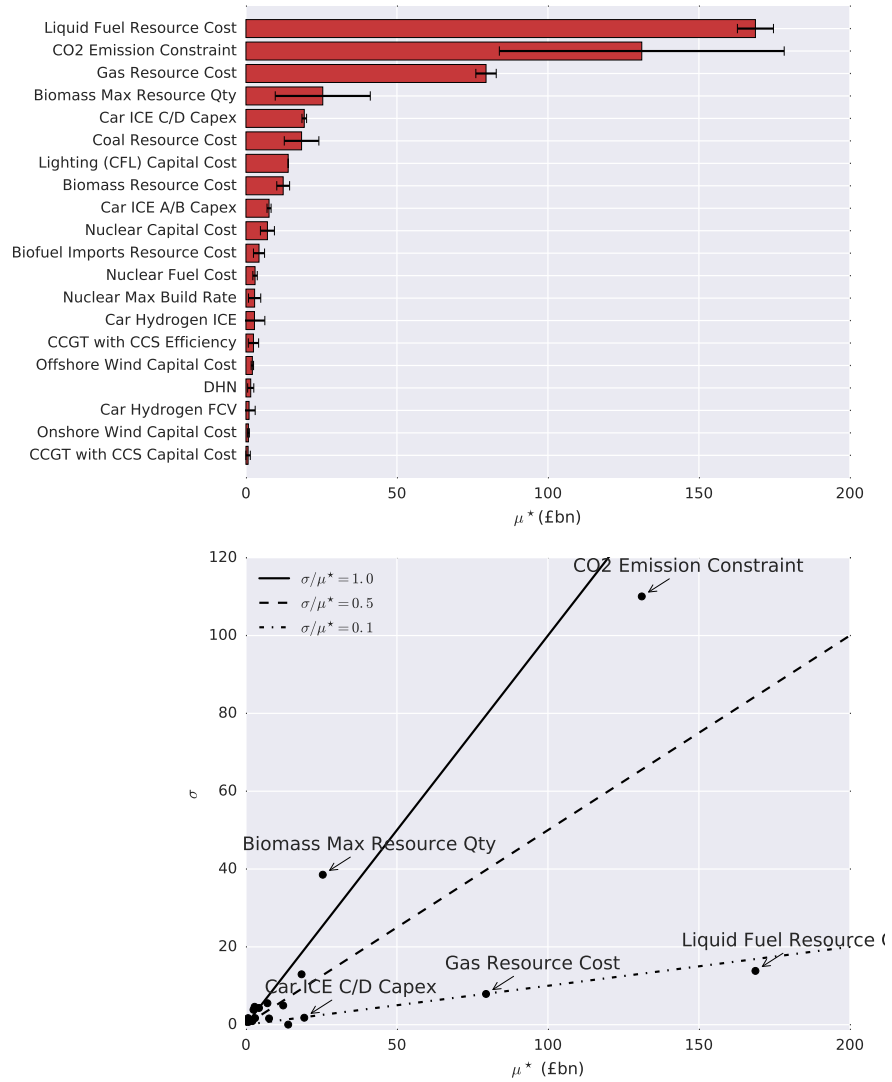


Figure 51: A more detailed analysis of the top 18 groups of parameters (of 44 tested) to which the ETI-Model is most sensitive.

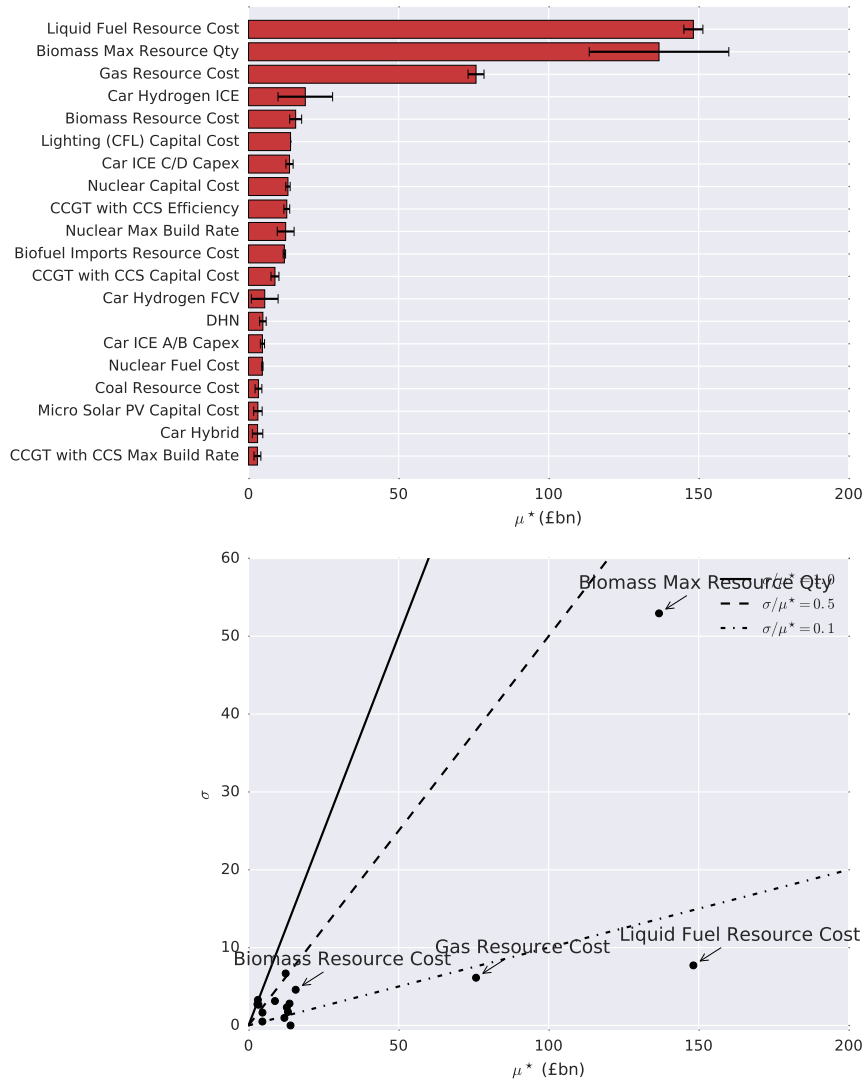
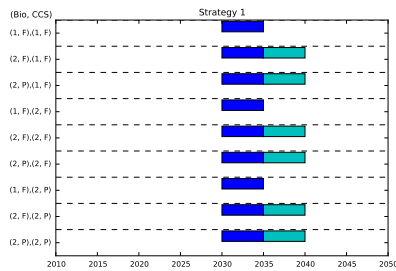


Figure 52: Results with carbon constraint set to 80% reduction by 2050

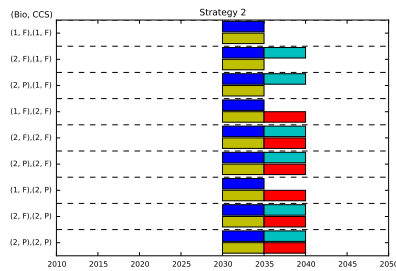
ADDITIONAL RESULTS: LEARNING MODEL

G.1 PREDICTED VALUES FOR THE REVENUE FUNCTION

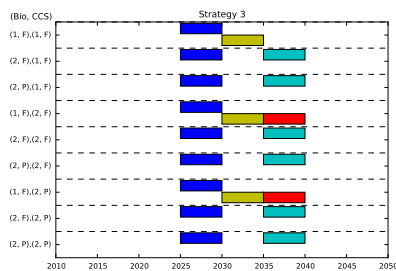
G.2 STRATEGIES FROM TWO-STAGE "BIG-PHARMA" PROJECTS



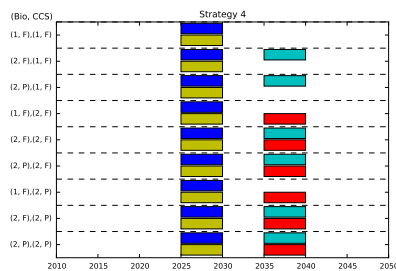
(a) Strategy 1



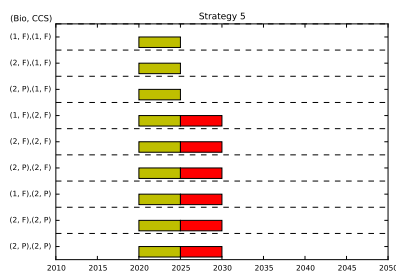
(b) Strategy 2



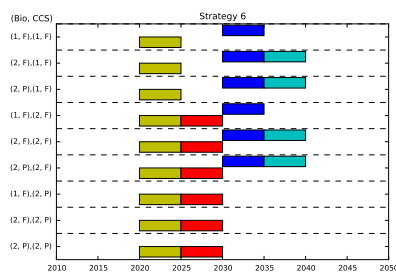
(c) Strategy 3



(d) Strategy 4



(e) Strategy 5



(f) Strategy 6

Figure 53: Strategies 1 to 6 from the two-stage big-pharma results in Section 6.3.1

Table 37: The modelled and predicted values for the revenue function

Simulation	Bio		CCS		Actual Revenue (£bn)	Predicted Revenue (£bn)
	2030	2040	2030	2040		
2	1	1	0	0	20.447	22.4106
3	0	1	0	0	18.001	20.4915
4	2	2	0	0	36.526	37.9620
5	1	2	0	0	35.424	36.0429
6	0	2	0	0	32.743	34.1238
7	0	0	1	1	27.596	25.3222
8	1	1	1	1	44.199	40.8736
9	0	1	1	1	41.903	38.9545
10	2	2	1	1	58.697	56.4250
11	1	2	1	1	NaN	54.5059
12	0	2	1	1	55.588	52.5868
13	0	0	0	1	20.291	21.2562
14	1	1	0	1	37.788	36.8076
15	0	1	0	1	35.087	34.8885
16	2	2	0	1	NaN	52.3590
17	1	2	0	1	51.811	50.4399
18	0	2	0	1	48.893	48.5208
19	0	0	2	2	41.939	43.7852
20	1	1	2	2	59.337	59.3366
21	0	1	2	2	57.354	57.4175
22	2	2	2	2	72.508	74.8880
23	1	2	2	2	71.705	72.9689
24	0	2	2	2	69.645	71.0498
25	0	0	1	2	39.152	39.7192
26	1	1	1	2	NaN	55.2706
27	0	1	1	2	54.277	53.3515
28	2	2	1	2	NaN	70.8220
29	1	2	1	2	NaN	68.9029
30	0	2	1	2	66.718	66.9838
31	0	0	0	2	34.327	35.6532
32	1	1	0	2	52.231	51.2046
33	0	1	0	2	49.258	49.2855
34	2	2	0	2	NaN	66.7560
35	1	2	0	2	65.066	64.8369
36	0	2	0	2	61.990	62.9178

Table 38: The revenue function for the "big-pharma" result set.

Simulation	Bio		CCS		Revenue (£bn)
	2030	2040	2030	2040	
2	1	1	0	0	0
3	0	1	0	0	0
4	2	2	0	0	36.526
5	1	2	0	0	32.743
6	0	2	0	0	32.743
7	0	0	1	1	0
8	1	1	1	1	0
9	0	1	1	1	0
10	2	2	1	1	36.526
11	1	2	1	1	32.743
12	0	2	1	1	32.743
13	0	0	0	1	0
14	1	1	0	1	0
15	0	1	0	1	0
16	2	2	0	1	36.526
17	1	2	0	1	32.743
18	0	2	0	1	32.743
19	0	0	2	2	41.939
20	1	1	2	2	41.939
21	0	1	2	2	41.939
22	2	2	2	2	72.508
23	1	2	2	2	69.645
24	0	2	2	2	69.645
25	0	0	1	2	34.327
26	1	1	1	2	34.327
27	0	1	1	2	34.327
28	2	2	1	2	66.756
29	1	2	1	2	66.756
30	0	2	1	2	61.990
31	0	0	0	2	34.327
32	1	1	0	2	34.327
33	0	1	0	2	34.327
34	2	2	0	2	66.756
35	1	2	0	2	61.990
36	0	2	0	2	61.990

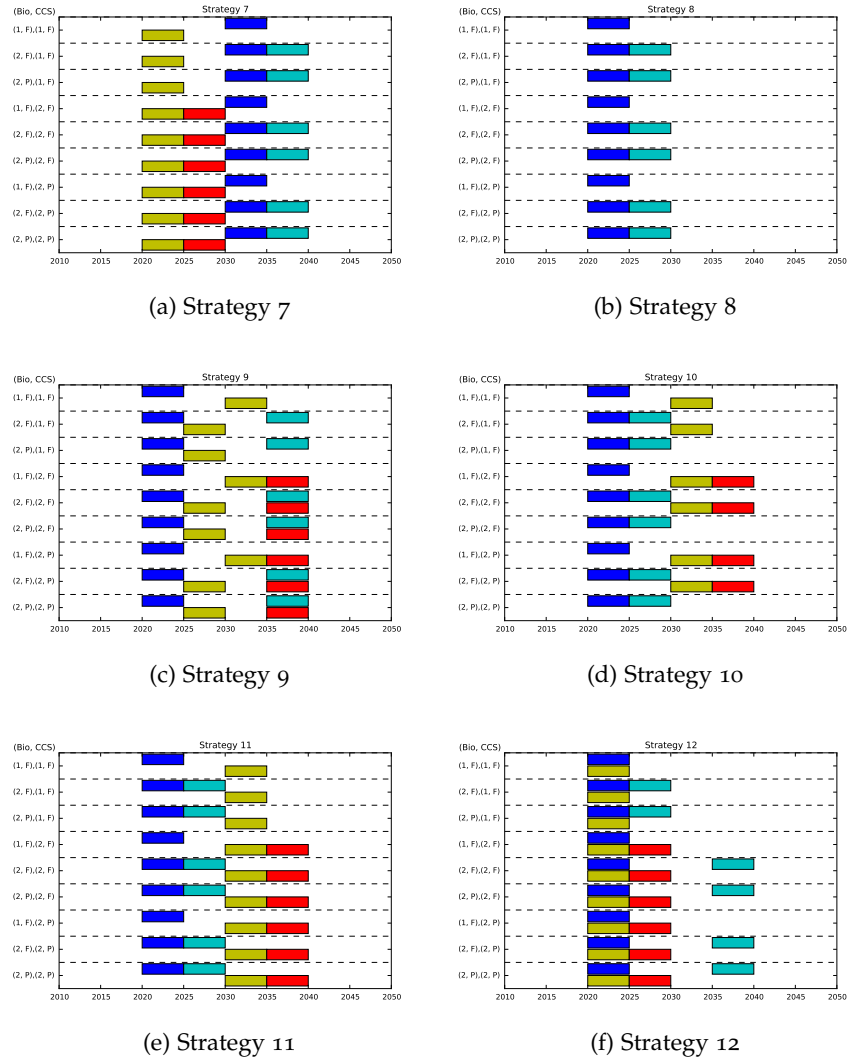


Figure 54: Strategies 7 to 12 from the two-stage big-pharma results in Section 6.3.1

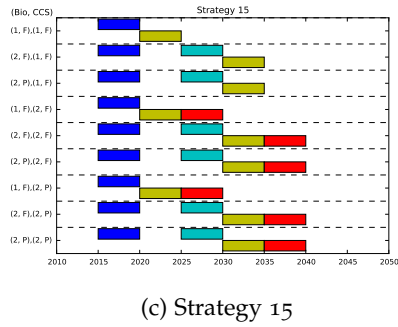
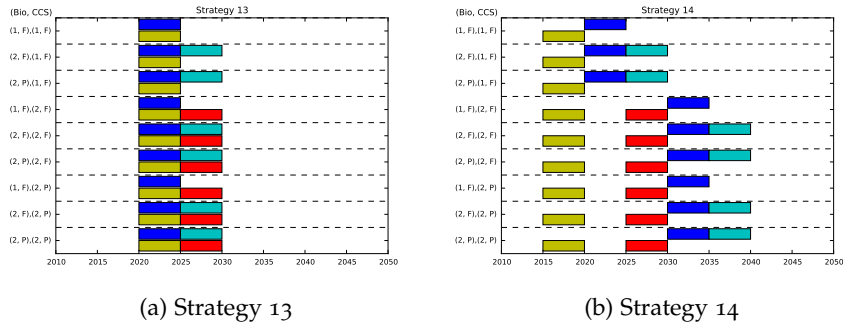
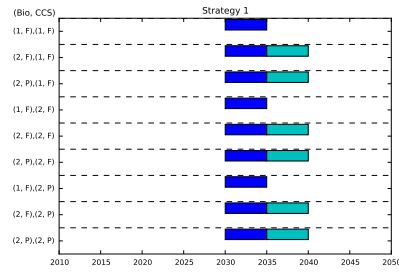
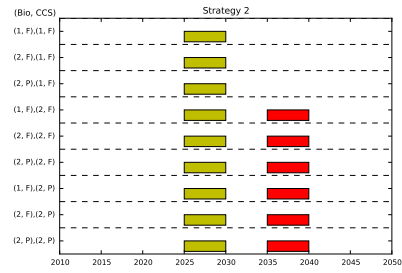


Figure 55: Strategies 13 to 15 from the two-stage big-pharma results in Section 6.3.1

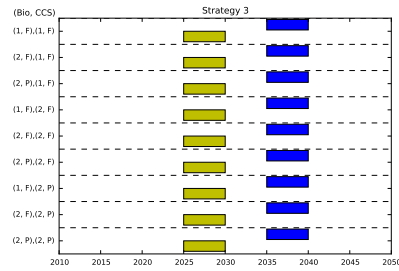
G.3 STRATEGIES FROM TWO-STAGE "PARTIAL-REVENUES" PROJECTS



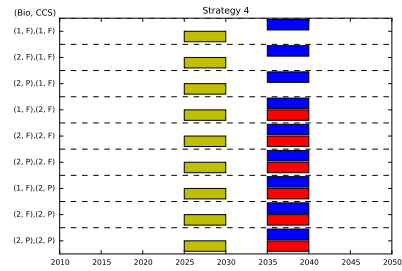
(a) Strategy 1



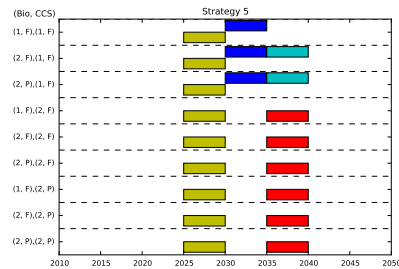
(b) Strategy 2



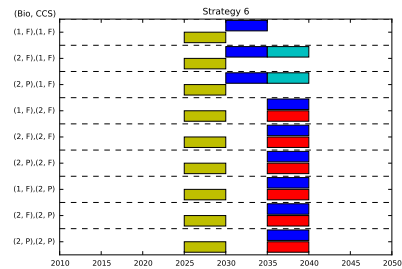
(c) Strategy 3



(d) Strategy 4

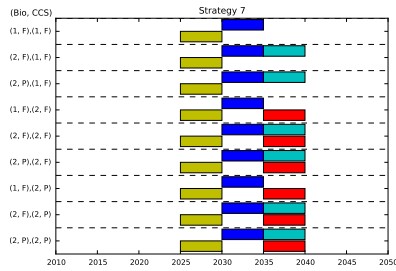


(e) Strategy 5

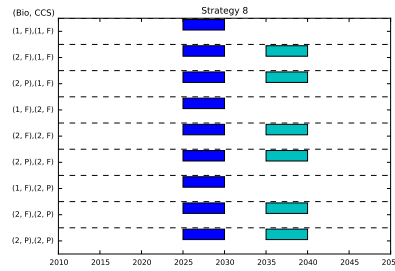


(f) Strategy 6

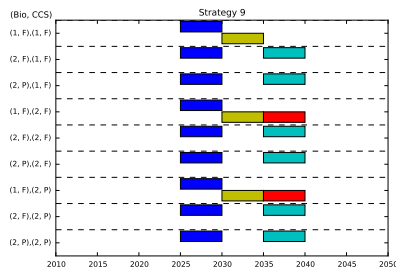
Figure 56: Strategies 1 to 6 from the two-stage partial projects in Section 6.3.2. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.



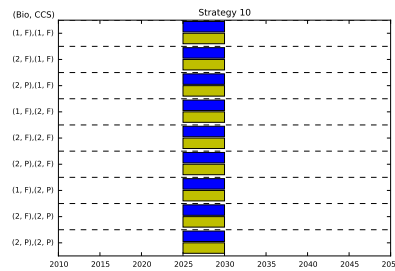
(a) Strategy 7



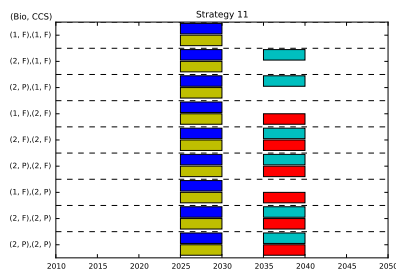
(b) Strategy 8



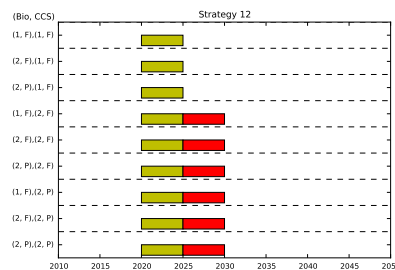
(c) Strategy 9



(d) Strategy 10



(e) Strategy 11



(f) Strategy 12

Figure 57: Strategies 7 to 12 from the two-stage partial projects in Section 6.3.2. Yellow and Red are CCS stages 1 and 2 respectively. Blue and Cyan are Bio Availability stages 1 and 2 respectively.

GLOSSARY

- adaptability** The ability of a strategy or policy to adapt to new, previously unknown outcomes. See also [flexibility](#), [diversity](#) and [resilience](#). [202](#)
- ADP** Approximate Dynamic Programming. A mathematical technique used for modelling sequential decision making under uncertainty. [30](#), [31](#)
- agency** The ability of a decision maker to act upon information, or the prospect of receiving new information. [38](#), [63](#)
- ambiguity** The case when probabilities are reasonably well known, but outcomes are not well defined (Stirling, [2007](#)). [2](#), [6](#), [227](#)
- BN** Bayesian Network. A probabilistic graphical model that represents the joint distribution of a set of random variables and enables the efficient calculation of marginal and conditional distribution functions (Poropudas et al., [2011](#)). [21](#)
- commercialisation** The process of taking an uncertain technology from an early or immature stage to the market. This could be treated as a parallel to the learning process of a [research project](#). [152](#), [175](#), [186](#), [189](#), [224](#)
- DDU** Decision-dependent uncertainty exists in two types. Type I is when a decision modifies the distribution of an uncertainty. Type II is when a decision results in the observation of an exogenous uncertainty. [3](#), [28](#), [223](#)
- decision variable** In an optimisation framework, decision variables are those whose values are chosen by the solver, i.e. as an output from the model. [3](#), [28](#), [29](#), [47](#)
- deterministic** A model in which the value of all input parameters are known with certainty. Compare with [stochastic](#). [223](#), [225](#)
- deterministic equivalent** The formulation of a stochastic scenario tree in which [NACs](#) are used to ensure that decision variables do not differ before the resolution of uncertainty in each branch of the scenario. The deterministic equivalent formulation allows large stochastic programming problems to be solved using commercial solvers. [103](#)
- diversity** A range of different or orthogonal strategies(Stirling, [2007](#), [2010](#)). [221](#)
- ESOM** Energy System Model. A linear programming optimisation model used to investigate energy transitions over long time horizons i.e. decades. Energy system models integrate energy, the environment, economics and technological factors

to provide cross-disciplinary insights. [xiv](#), [1](#), [5–7](#), [9–11](#), [13–15](#), [20](#), [22](#), [31](#), [32](#), [35](#), [37](#), [38](#), [41](#), [50–52](#), [64](#), [68](#), [74](#), [90](#), [168](#), [173](#), [176](#), [177](#), [180](#), [181](#), [188](#)

ETL Endogenous technological learning. The internal (to a model) calculation of the reduction in technology costs as a function of cumulative installed capacity. [34](#), [35](#), [222](#)

evolvability The capacity of a policy for adaptive evolution, for example a policy to act as the basis of a new policy to meet new needs or attain new levels of capability (McManus et al., [2006](#)). [202](#)

flexibility The ability of a strategy or policy to be modified to cover aspects for which it was not originally intended (McManus et al., [2006](#)). Common terminology in [real option](#) to refer to the ability of a decision maker to defer a decision until better information arrives. This flexibility is akin to [VOI](#). [202](#), [221](#), [224](#)

global sensitivity analysis When a [sensitivity analysis](#) is performed over the entire input space of the model, e.g. for a significant proportion of the parameters, and for each combination of parameters over the full range of possible values. [7](#), [8](#), [10](#), [38](#), [42](#), [53](#), [58](#), [67](#), [82](#), [89](#), [92](#), [167–170](#), [177–181](#), [185](#), [223](#), [226](#)

hurdle rate Technology specific discount rate, often used to represent non-financial costs associated with a particular technology. [50](#)

ignorance A lack of knowledge regarding both the consequences and likelihood of an event, or system. Note that this extends to the very ability to define what a problem is (Stirling, [2007](#)). [2](#), [6](#), [33](#), [227](#)

learning Reducing one's epistemic uncertainty or ignorance about the future. Different to the learning in [ETL](#). Learning in this context is related to [VOI](#). Webster et al., ([2008](#)) states learning is the "... idealized [sic] process of narrowing uncertainty..." Oppenheimer et al., ([2008](#)) define learning as "... the change in any aspect of the uncertainty of an outcome occurring as a result of theory development, modeling [sic], observation, or experiment". [3](#), [32](#), [176](#), [224](#)

negative learning The converse of [progressive learning](#), negative learning occurs when successive learning events result in a less accurate depiction of the uncertainty i.e. further from the true value. [32](#)

progressive learning Learning is progressive if successive learning events result in more accurate depiction of the uncertainty i.e. closer to the true value. [222](#)

level One of an even number of equally sized components that derive from the division of the input space of a parameter into discrete parts. For example, four levels of the uniform distribution [0, 1] gives [0.00, 0.25], [0.25, 0.50], [0.50, 0.75] and [0.75, 1.00]. These are used in the [global sensitivity analysis](#) Method of Morris. [25](#), [53](#), [54](#), [58](#), [59](#), [67](#), [70](#), [193](#), [194](#)

metrics Metrics are...

EVBI The value of information gained through a reduction in uncertainty, when the new information is compared against a baseline in a model (as defined in Baker et al., [2011c](#)). [37](#)

EVIU The expected value of including uncertainty. [36](#)

EVPI The Expected Value of Perfect Information in a two-stage stochastic model is the difference in objective function cost terms of solutions with perfect information, and those in which uncertainty is modelled. [36](#), [37](#)

VOI Value of Information. [36](#), [222](#)

VSS A measure to quantify the advantage of using a [stochastic programming](#) approach over a [deterministic](#) one. [36](#)

Monte Carlo A Monte Carlo simulation is a technique for uncertainty propagation in which the inputs for successive runs of a deterministic model are sampled from a joint distribution of input parameters. [21](#), [176](#), [226](#), [227](#)

NAC Non-anticipativity constraints are imposed in the deterministic equivalent of a stochastic programming problem to ensure that decision variables cannot anticipate uncertain information before the uncertainty is resolved. In the case of exogenous uncertainties, the constraints are fixed. When uncertainties are [decision-dependent](#), then a mechanism for adjusting when to impose the constraints must be introduced, increasing the complexity of the formulation. [28](#), [30](#), [103](#), [106](#), [123](#), [152](#), [169](#), [221](#), [225](#)

outcome The enumeration of the results of [research project stages](#) in the learning model (pass/fail) *independent of when or whether investment in the project occurs*. The outcomes are associated with probabilities, derived from the probability of success of each [stage](#) in a [project](#). [xv](#), [xvi](#), [106–108](#), [111](#), [118](#), [119](#), [123](#), [125](#), [131](#), [132](#), [134](#), [152](#), [154](#), [188](#), [190](#), [224](#), [225](#)

parameterisation describe or represent a part of a system in a model using parameters, instead of explicitly modelling the underlying processes. [21](#), [178](#), [203](#)

project . [123](#), [152](#), [223](#), [225](#), *see* [research project](#)

R&D A process of generating knowledge about a subject or technology. A successful R&D process is normally succeed by commercialisation of a technology. [xv](#), [xvi](#), [107](#), [197](#)

- real option** The application of financial options theory to real problems, as outlined in Dixit et al., 1994. Real options models the flexibility of a decision maker to wait for better information, and can value this option to wait. 222
- research project** A sequence of one or more individual stages whose outcomes are linked to discrete values of uncertain parameters in the ESME model. The mechanism by which uncertainty is resolved within a research project is left intentionally vague in this thesis, but could include learning or commercialisation. 2, 3, 5, 10, 11, 97–101, 103, 104, 106–110, 113–116, 118, 120, 122, 123, 125–130, 133–139, 141, 145, 147, 150, 152, 154, 159, 167, 170, 171, 173–176, 180, 182–190, 221, 223–225
- resilience** A strategy is resilient if it has the capacity to recover quickly from (previously unforeseen) difficulties. 221
- revenue function** A vector of costs associated with leaves of the scenario tree built from the outcomes of research projects from the ESME model. An important input to the learning model, the values associated with the research project outcomes define the results from the model in conjunction with the probability of success of a stage. 97, 106, 111–113, 119, 125, 126, 129–131, 135, 138, 141, 147, 169, 175, 182, 183, 185, 186
- risk** A confusing array of definitions of risk exist. In the quantitative risk-assessment field, ‘risk’ refers to the product of the probability and cost associated with an undesirable event. 1, 2, 202, 227
- Knightian Risk** Uncertainty in which the probabilities associated with an undesirable event are so well defined and known that a quantitative analysis is possible e.g. using expected-cost or -utility to support decision making. 58, 201, 202, 204
- robustness** A strategy is robust if it performs within a tolerable threshold of performance over a range of scenarios (Lempert et al., 2007) or unexpectedly adverse environments (McManus et al., 2006). Robustness can be quantified through modelling, using Maxi-min or Minimax-regret objective functions in an optimisation model. Lempert et al., (2007) shows that the existence of robust strategies depends on both the degree of uncertainty and the flexibility available to the decision maker. 202
- scenario** The combinations of successful research projects and stages for each key year in the ESME model. The scenarios define the steps that make up the revenue function from the ESME model (outlined in 5.2.2) and quantify the value of each successful research investment strategy. 110–113, 125, 224, 225
- scenario analysis** An internally consistent, plausible vision of the future, often packaged into a two-by-two matrix. Initially developed for military and strategic purposes in the 50s and

- 60s, and later adopted by Shell and others increasingly as a response to the failure of forecasting and projections. 9, 13, 15, 21, 41, 48, 50, 177–179
- scenario tree** The description of a discrete probability distribution, branching from the trunk (a known starting point) to many leaf nodes (an uncertain future). The scenario tree can be described mathematically using parameter values indexed by stochastic scenario (leaf node), while NACs are used to ensure the structure of the tree. 28–30, 36, 37, 98, 103, 106–109, 117, 123, 132, 152, 172
- SDP** Stochastic dynamic programming. A general approach to modelling sequential decision making under uncertainty, where a state is updated by actions in response to uncertain information. An optimal policy (sequence of actions) is sought that minimises/maximises cost/utility. 13, 28, 30, 225
- sensitivity analysis** Computation of effect of changes in inputs on model outputs. Techniques of varying comprehensiveness are documented in Saltelli et al., (2000) spanning computationally benign methods for parameter screening, through to computation of main effects, to computationally expensive calculation of 2nd- and higher-order interactions. 11, 15, 22–27, 36, 38, 178, 222, 226, 227
- stage** An individual phase of a sequential [research project](#). A stage has a binary success/fail [outcome](#) and has an associated probability. Contrast with [project](#), [outcome](#), [scenario](#). xiv, xvi, 10, 98–111, 113–123, 125, 127–138, 141–147, 149–152, 154, 155, 159, 167, 170–173, 175, 176, 182–185, 187, 189, 191, 223, 224
- stochastic** Random. In relation to models, a stochastic model is one in which some or all input parameters take a random value according to a defined probability distribution. Compare with [deterministic](#). 221
- stochastic programming** A technique in operations research, in which the objective function computes the expected cost (or equivalent) given two or more probability weighted future states of the world. Alternatively, a special case of SDP where the probability distributions representing uncertainties are known before solving the model. 13, 34, 36, 223, 226
- strategy** The collection of investment decisions from the learning model representing the optimal response to ongoing resolution of uncertainty over time. A strategy is an output from the learning model. 1, 125, 126, 132, 134–140, 145, 147, 150
- stride** A term borrowed from computer science, relating to the vector form of a matrix (e.g. to store a multi-dimensional matrix as a vector in memory). The stride refers to the size of one dimension of the original matrix, which when in vector form,

allows an index counter to jump to the next relevant entry in the vector. [xvi](#), [107](#)

trajectory In the [global sensitivity analysis](#) technique, Method of Morris, a “path” of individual parameter steps over a finite number of levels. Each trajectory contains $k + 1$ entries, where k is the number of parameters. Each trajectory is randomly generated from a different starting position. N trajectories are used to complete an analysis, where N is as large as feasible. Each trajectory provides one sample for each parameter, from which the results metrics may be computed. [25](#), [52–58](#), [60](#), [61](#), [67](#), [68](#), [70](#)

UA Uncertainty Analysis: Compare the importance of input uncertainties in terms of relative contributions to uncertainty in the outputs. See also [UP](#), [sensitivity analysis](#), [Monte Carlo](#). [22](#), [23](#), [27](#), [227](#)

uncertainty The lack of certainty about an event, value or definition. Or, according to Walker et al., (2003) “. . . any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system. . .” There are numerous, often conflicting definitions of uncertainty, so rather than present an overview of the full range of perspectives, I define a subset used consistently in this document. [1](#), [2](#)

aleatory uncertainty natural randomness, uncertainty that cannot be reduced further. [2](#), [8](#), [63](#), [203](#)

dynamic uncertainty The type of uncertainty which can change over time. See [epistemic uncertainty](#). [2–4](#), [6–8](#), [13](#), [28](#), [29](#), [31](#), [33](#), [38](#)

endogenous uncertainty In decision dependent uncertainty, either decisions can affect the probability distribution of the uncertainties themselves. Or decisions can affect the timing of the resolution of uncertainty. [3](#), [29](#), [32](#), [99](#), [103](#)

epistemic uncertainty epistemic refers to knowledge, and thus epistemic uncertainty is that associated with lack of knowledge or learning. An uncertainty can be classified as epistemic if it can be reduced through, for example, waiting, learning or conducting research. [2](#), [3](#), [32](#), [63](#), [203](#), [226](#)

exogenous uncertainty In [stochastic programming](#), the uncertainty is resolved outside the model and at fixed time-periods regardless of the decisions taken. [3](#), [4](#), [29](#), [35](#), [103](#)

Keynesian uncertainty Equivalent to [Knightian uncertainty](#). [227](#)

Knightian uncertainty Uncertainty for which a probability distribution can not be estimated, or approximated. This concept is named after Frank Knight and was raised in his seminal work (see Knight, 1921). Equivalent to [Keynesian uncertainty](#). Contrast with [ambiguity](#), [ignorance](#) and [risk](#). [2](#), [6](#), [41](#), [202](#), [226](#)

- parametric uncertainty** Uncertainty associated with the inputs (parameters) to a model. [8](#), [10](#), [14](#), [15](#), [21](#), [37](#)
- structural uncertainty** Uncertainty associated with the mathematical formulation of a model. [10](#), [13](#), [14](#), [37](#)
- UP** Uncertainty Propagation: Calculate the uncertainty in model outputs induced by uncertain inputs. See also [UA](#), [sensitivity analysis](#), [Monte Carlo](#). [22](#), [23](#), [226](#)

BIBLIOGRAPHY

- Agnew, Malcolm, Leo Schratzenholzer and Alfred Voss (1979). 'A model for energy supply systems alternatives and their general environmental impact'. Lexenbug, Austria.
- Akimoto, Keigo et al. (2010). 'Estimates of GHG emission reduction potential by country, sector, and cost'. In: *Energy Policy* 38.7, pp. 3384–3393. ISSN: 03014215. DOI: [10.1016/j.enpol.2010.02.012](https://doi.org/10.1016/j.enpol.2010.02.012). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421510000935>.
- Anadon, Laura Diaz (2012). 'Missions-oriented RD&D institutions in energy between 2000 and 2010: A comparative analysis of China, the United Kingdom, and the United States'. In: *Research Policy* 41.10, pp. 1742–1756. ISSN: 00487333. DOI: [10.1016/j.respol.2012.02.015](https://doi.org/10.1016/j.respol.2012.02.015). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0048733312002211>.
- Anadon, Laura Diaz et al. (2011). *Transforming U.S. Energy Innovation*. Tech. rep. November. Belfer Center for Science and International Affairs, Harvard Kennedy School, p. 340.
- Anadon, Laura Diaz et al. (2012). 'Expert judgments about RD&D and the future of nuclear energy'. In: *Environmental Science and Technology* 46.21, pp. 11497–11504. ISSN: 0013936X. DOI: [10.1021/es300612c](https://doi.org/10.1021/es300612c). URL: <http://www.ncbi.nlm.nih.gov/pubmed/23002786>.
- Anadon, Laura Diaz, G. Chan and A Lee (2014). *Expanding and better targeting U.S. investment in energy innovation: An analytical approach*, pp. 36–80. ISBN: 9781107043718. DOI: [10.1017/CB09781107338890.002](https://doi.org/10.1017/CB09781107338890.002).
- Anadon, Laura Diaz et al. (2016). 'Expert views - and disagreements - about the potential of energy technology R&D'. In: *Climatic Change*, pp. 1–15. ISSN: 15731480. DOI: [10.1007/s10584-016-1626-0](https://doi.org/10.1007/s10584-016-1626-0). URL: <http://dx.doi.org/10.1007/s10584-016-1626-0>.
- Anandarajah, G and Neil Strachan (2010). 'Interactions and implications of renewable and climate change policy on UK energy scenarios'. In: *Energy Policy* 38.11, pp. 6724–6735. ISSN: 03014215. DOI: [10.1016/j.enpol.2010.06.042](https://doi.org/10.1016/j.enpol.2010.06.042). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421510005069>.
- Anderson, Barry et al. (2014). 'Uncertainty in climate change modeling: can global sensitivity analysis be of help?' In: *Risk analysis : an official publication of the Society for Risk Analysis* 34.2, pp. 271–93. ISSN: 1539-6924. DOI: [10.1111/risa.12117](https://doi.org/10.1111/risa.12117). URL: <http://www.ncbi.nlm.nih.gov/pubmed/24111855>.
- Archer, G. E. B., Andrea Saltelli and I. M. Sobol' (1997). 'Sensitivity measures, anova-like Techniques and the use of bootstrap'. In: *Journal of Statistical Computation and Simulation* 58.2, pp. 99–120.

- ISSN: 0094-9655. DOI: [10.1080/00949659708811825](https://doi.org/10.1080/00949659708811825). URL: <http://www.tandfonline.com/doi/abs/10.1080/00949659708811825>.
- Arrow, Kenneth J (1962). 'The Economic Learning Implications of by Doing'. In: *The Review of Economic Studies* 29.3, pp. 155–173. URL: <http://www.jstor.org/stable/2295952>.
- Asselt, M.B.A. van (2000). *Perspectives on uncertainty and risk: the PRIMA approach to decision support*. Kluwer Academic Publishers. ISBN: 9780792366560.
- Asselt, M.B.A. van and Jan Rotmans (2002). 'Uncertainty in integrated assessment modelling: From Positivism to Pluralism'. In: *Climatic Change* 54.1-2, pp. 75–105. DOI: [10.1023/A:1015783803445](https://doi.org/10.1023/A:1015783803445). URL: <http://www.springerlink.com/index/yd1q6q9x9tftp6yxj.pdf>.
- Avenhaus, Rudolf et al. (1980). 'Handling uncertainties in Linear Programming Models'. Lexenbug, Austria.
- Babonneau, Frédéric et al. (2011a). 'Combining Stochastic Optimization and Monte Carlo Simulation to Deal with Uncertainties in Climate Policy Assessment'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 51–76. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9275-1](https://doi.org/10.1007/s10666-011-9275-1). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9275-1>.
- Babonneau, Frédéric et al. (2011b). 'Energy Security: A Robust Optimization Approach to Design a Robust European Energy Supply via TIAM-WORLD'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 19–37. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9273-3](https://doi.org/10.1007/s10666-011-9273-3). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9273-3>.
- Bajželj, Bojana et al. (2014). 'Importance of food-demand management for climate mitigation'. In: *Nature Climate Change* advance on. ISSN: 1758-678X. DOI: [10.1038/nclimate2353](https://doi.org/10.1038/nclimate2353). URL: <http://www.nature.com/doifinder/10.1038/nclimate2353http://dx.doi.org/10.1038/nclimate2353http://www.nature.com/nclimate/journal/vaop/ncurrent/abs/nclimate2353.html{\#}supplementary-information>.
- Baker, Erin, Leon Clarke and John Weyant (2006). 'Optimal Technology R&D in the Face of Climate Uncertainty'. In: *Climatic Change* 78.1, pp. 157–179. ISSN: 0165-0009. DOI: [10.1007/s10584-006-9092-8](https://doi.org/10.1007/s10584-006-9092-8). URL: <http://www.springerlink.com/content/q74p624268913p4g/http://link.springer.com/10.1007/s10584-006-9092-8>.
- Baker, Erin and Kwame Adu-Bonnah (2008). 'Investment in risky R&D programs in the face of climate uncertainty'. In: *Energy Economics* 30.2, pp. 465–486. ISSN: 01409883. DOI: [10.1016/j.eneco.2006.10.003](https://doi.org/10.1016/j.eneco.2006.10.003).
- Baker, Erin, Haewon Chon and Jeffrey Keisler (2009). 'Carbon capture and storage: combining economic analysis with expert elicitations to inform climate policy'. In: *Climatic Change* 96.3, pp. 379–408.

- ISSN: 0165-0009. DOI: [10.1007/s10584-009-9634-y](https://doi.org/10.1007/s10584-009-9634-y). URL: <http://www.springerlink.com/index/10.1007/s10584-009-9634-y>
<http://link.springer.com/10.1007/s10584-009-9634-y>.
- Baker, Erin, Haewon Chon and Jeffrey M. Keisler (2010). 'Battery technology for electric and hybrid vehicles: Expert views about prospects for advancement'. In: *Technological Forecasting and Social Change* 77.7, pp. 1139–1146. ISSN: 00401625. DOI: [10.1016/j.techfore.2010.02.005](https://doi.org/10.1016/j.techfore.2010.02.005). URL: <http://linkinghub.elsevier.com/retrieve/pii/S004016251000034X>.
- Baker, Erin and S.b Solak (2011a). 'Climate change and optimal energy technology R&D policy'. In: *European Journal of Operational Research* 213.2, pp. 442–454. URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-79955880661&partnerID=40&md5=4b60a21fa93456bf7a9108ead9818f4e>.
- Baker, Erin and Senay Solak (2011b). 'Climate change and optimal energy technology R&D policy'. In: *European Journal of Operational Research* 213.2, pp. 442–454. ISSN: 03772217. DOI: [10.1016/j.ejor.2011.03.046](https://doi.org/10.1016/j.ejor.2011.03.046). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0377221711003080>.
- Baker, Erin and Yiming Peng (2011c). 'The Value of Better Information on Technology R&D Programs in Response to Climate Change'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 107–121. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9278-y](https://doi.org/10.1007/s10666-011-9278-y). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9278-y>.
- Baker, Erin, Olaitan Olaleye and Lara Aleluia Reis (2015). 'Decision frameworks and the investment in R&D'. In: *Energy Policy* 80, pp. 275–285. ISSN: 03014215. DOI: [10.1016/j.enpol.2015.01.027](https://doi.org/10.1016/j.enpol.2015.01.027). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421515000361>.
- Barreto, Leonardo and Socrates Kypreos (2004). 'Emissions trading and technology deployment in an energy-systems bottom-up model with technology learning'. In: *European Journal of Operational Research* 158.1, pp. 243–261. ISSN: 03772217. DOI: [10.1016/S0377-2217\(03\)00350-3](https://doi.org/10.1016/S0377-2217(03)00350-3). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0377221703003503>.
- Barron, Robert and Haewon McJeon (2015). 'The differential impact of low-carbon technologies on climate change mitigation cost under a range of socioeconomic and climate policy scenarios'. In: *Energy Policy* 80, pp. 264–274. ISSN: 03014215. DOI: [10.1016/j.enpol.2015.01.038](https://doi.org/10.1016/j.enpol.2015.01.038). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421515000579>.
- Bazilian, Morgan et al. (2011). 'Considering the energy, water and food nexus: Towards an integrated modelling approach'. In: *Energy Policy*, pp. 1–11. ISSN: 03014215. DOI: [10.1016/j.enpol.2011.09.039](https://doi.org/10.1016/j.enpol.2011.09.039). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421511007282>.

- Bellman, Richard (1952). 'On the Theory of Dynamic Programming'. In: *Proceedings of the National Academy of Sciences of the United States of America* 38.8, pp. 716–719. ISSN: 0027-8424.
- Berger, C (1990). 'Modelling energy use in the mineral processing industries of Ontario with MARKAL-Ontario'. In: *Energy* 15.9, pp. 741–758. ISSN: 03605442. DOI: [10.1016/0360-5442\(90\)90113-G](https://doi.org/10.1016/0360-5442(90)90113-G). URL: <http://linkinghub.elsevier.com/retrieve/pii/036054429090113G>.
- Bertsekas, Dimitri P. (2005). *Dynamic Programming and Optimal Control*. 3rd Ed. Nashua, NH, USA: Athena Scientific. ISBN: 1886529086.
- Beugin, Dale and Mark Jaccard (2011). 'Statistical Simulation to Estimate Uncertain Behavioral Parameters of Hybrid Energy-Economy Models'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 77–90. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9276-0](https://doi.org/10.1007/s10666-011-9276-0). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9276-0>.
- Bosetti, Valentina and Massimo Tavoni (2009). 'Uncertain R&D, backstop technology and GHGs stabilization'. In: *Energy Economics* 31.SUPPL. 1, S18–S26. ISSN: 01409883. DOI: [10.1016/j.eneco.2008.03.002](https://doi.org/10.1016/j.eneco.2008.03.002). URL: <http://dx.doi.org/10.1016/j.eneco.2008.03.002>.
- Bosetti, Valentina et al. (2015). 'Sensitivity to energy technology costs: A multi-model comparison analysis'. In: *Energy Policy* 80, pp. 244–263. ISSN: 03014215. DOI: [10.1016/j.enpol.2014.12.012](https://doi.org/10.1016/j.enpol.2014.12.012). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421514006776>.
- Branger, Frédéric et al. (2015). 'Global sensitivity analysis of an energy-economy model of the residential building sector'. In: *Environmental Modelling & Software* 70, pp. 45–54. ISSN: 13648152. DOI: [10.1016/j.envsoft.2015.03.021](https://doi.org/10.1016/j.envsoft.2015.03.021). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364815215001097>.
- Bréchet, Thierry et al. (2011). 'The Benefits of Cooperation Under Uncertainty: the Case of Climate Change'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 149–162. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9281-3](https://doi.org/10.1007/s10666-011-9281-3). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9281-3>.
- Brennan, Alan et al. (2007). 'Calculating partial expected value of perfect information via Monte Carlo sampling algorithms.' In: *Medical Decision Making* 27.4, pp. 448–470. URL: <http://www.ncbi.nlm.nih.gov/pubmed/17761960>.
- Broek, Machteld van den et al. (2009). 'An integrated GIS-MARKAL toolbox for designing a CO₂ infrastructure network in the Netherlands'. In: *Energy Procedia* 1.1, pp. 4071–4078. ISSN: 18766102. DOI: [10.1016/j.egypro.2009.02.214](https://doi.org/10.1016/j.egypro.2009.02.214). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1876610209008571>.
- Broek, Machteld van den et al. (2010). 'Feasibility of storing CO₂ in the Utsira formation as part of a long term Dutch CCS strategy An evaluation based on a GIS/MARKAL toolbox'. In: *International Journal of Greenhouse Gas Control* 4.2, pp. 351–366. ISSN: 17505836.

- DOI: [10.1016/j.ijggc.2009.09.002](https://doi.org/10.1016/j.ijggc.2009.09.002). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1750583609000905>.
- Butler, Martha P. et al. (2014). 'Identifying parametric controls and dependencies in integrated assessment models using global sensitivity analysis'. In: *Environmental Modelling & Software* 59, pp. 10–29. ISSN: 13648152. DOI: [10.1016/j.envsoft.2014.05.001](https://doi.org/10.1016/j.envsoft.2014.05.001). URL: <http://www.sciencedirect.com/science/article/pii/S1364815214001327>.
- Campolongo, F and R. Braddock (1999). 'Sensitivity analysis of the IMAGE greenhouse model'. In: *Environmental Modelling and Software* 14.4, pp. 275–282. ISSN: 13648152. DOI: [10.1016/S1364-8152\(98\)00079-6](https://doi.org/10.1016/S1364-8152(98)00079-6).
- Campolongo, F, Jessica Cariboni and Andrea Saltelli (2007). 'An effective screening design for sensitivity analysis of large models'. In: *Environmental Modelling & Software* 22.10, pp. 1509–1518. ISSN: 13648152. DOI: [10.1016/j.envsoft.2006.10.004](https://doi.org/10.1016/j.envsoft.2006.10.004). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364815206002805>.
- Campolongo, F, Andrea Saltelli and Jessica Cariboni (2011). 'From screening to quantitative sensitivity analysis. A unified approach'. In: *Computer Physics Communications* 182.4, pp. 978–988. ISSN: 00104655. DOI: [10.1016/j.cpc.2010.12.039](https://doi.org/10.1016/j.cpc.2010.12.039). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0010465510005321>.
- Chen, W (2005). 'The costs of mitigating carbon emissions in China: findings from China MARKAL-MACRO modeling'. In: *Energy Policy* 33.7, pp. 885–896. ISSN: 03014215. DOI: [10.1016/j.enpol.2003.10.012](https://doi.org/10.1016/j.enpol.2003.10.012). URL: <http://linkinghub.elsevier.com/retrieve/pii/S030142150300315X>.
- Chen, W et al. (2007). 'Carbon emission control strategies for China: A comparative study with partial and general equilibrium versions of the China MARKAL model'. In: *Energy* 32.1, pp. 59–72. ISSN: 03605442. DOI: [10.1016/j.energy.2006.01.018](https://doi.org/10.1016/j.energy.2006.01.018). URL: <http://linkinghub.elsevier.com/retrieve/pii/S036054420600034X>.
- Childs, Paul D. and Alexander J. Triantis (1999). 'Dynamic R&D Investment Policies'. In: *Management Science* 45.10, pp. 1359–1377. ISSN: 0025-1909. DOI: [10.1287/mnsc.45.10.1359](https://doi.org/10.1287/mnsc.45.10.1359). URL: <http://pubsonline.informs.org/doi/abs/10.1287/mnsc.45.10.1359>.
- Cian, Enrica and Massimo Tavoni (2011). 'Mitigation Portfolio and Policy Instruments When Hedging Against Climate Policy and Technology Uncertainty'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 123–136. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9279-x](https://doi.org/10.1007/s10666-011-9279-x). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9279-x>.
- Cinar, Didem and Gulgun Kayakutlu (2010). 'Scenario analysis using Bayesian networks: A case study in energy sector'. In: *Knowledge-Based Systems* 23.3, pp. 267–276. ISSN: 09507051. DOI: [10.1016/j.kbs.2010.05.001](https://doi.org/10.1016/j.kbs.2010.05.001).

- knosys.2010.01.009. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0950705110000110>.
- Colvin, Matthew and Christos T. Maravelias (2008). 'A stochastic programming approach for clinical trial planning in new drug development'. In: *Computers & Chemical Engineering* 32.11, pp. 2626–2642. ISSN: 00981354. DOI: 10.1016/j.compchemeng.2007.11.010. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0098135407002876>.
- Colvin, Matthew and Christos T. Maravelias (2010). 'Modeling methods and a branch and cut algorithm for pharmaceutical clinical trial planning using stochastic programming'. In: *European Journal of Operational Research* 203.1, pp. 205–215. ISSN: 03772217. DOI: 10.1016/j.ejor.2009.07.022. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0377221709005244>.
- Condevaux-Lanloy, Christian and Emmanuel Fragniere (1996). 'A stochastic programming model for energy/environment choices under uncertainty'. In: *INTERNATIONAL JOURNAL OF ENVIRONMENT AND POLLUTION* 6.4-6, pp. 587–603. ISSN: 0957-4352. URL: http://apps.isiknowledge.com/full{_}record.do?product=UA{\&}search{_}mode=GeneralSearch{\&}qid=1{\&}SID=R1JdPN6ikDkNjJoN2@I{\&}page=1{\&}doc=1{\&}colname=WOS.
- Condevaux-Lanloy, Christian and Emmanuel Fragniere (2000). 'An approach to deal with uncertainty in energy and environmental planning: the MARKAL case'. In: *Environmental Modeling and Assessment* 5, pp. 145–155. URL: <http://www.springerlink.com/index/t1457t134241158p.pdf>.
- Conejo, Antonio J., Miguel Carrión and Juan M. Morales (2010). *Decision Making Under Uncertainty in Electricity Markets*. Vol. 153. International Series in Operations Research & Management Science. Boston, MA: Springer US. ISBN: 978-1-4419-7420-4. DOI: 10.1007/978-1-4419-7421-1. URL: <http://metro-natshar-31-71.brain.net.pk/articles/1441964711.pdf>http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=1709323<http://www.springerlink.com/index/10.1007/978-1-4419-7421-1>.
- Contaldi, M, F Gracceva and GC Tosato (2007). 'Evaluation of green-certificates policies using the MARKAL-MACRO-Italy model'. In: *Energy Policy* 35.2, pp. 797–808. ISSN: 03014215. DOI: 10.1016/j.enpol.2006.03.011. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421506001157>.
- Contaldi, M, F Gracceva and a Mattucci (2008). 'Hydrogen perspectives in Italy: Analysis of possible deployment scenarios'. In: *International Journal of Hydrogen Energy* 33.6, pp. 1630–1642. ISSN: 03603199. DOI: 10.1016/j.ijhydene.2007.12.035. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360319907007434>.
- Contreras, a, E Guervos and F Posso (2009). 'Market penetration analysis of the use of hydrogen in the road transport sector of the

- Madrid region, using MARKAL'. In: *International Journal of Hydrogen Energy* 34.1, pp. 13–20. ISSN: 03603199. DOI: [10.1016/j.ijhydene.2008.10.031](https://doi.org/10.1016/j.ijhydene.2008.10.031). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360319908013086>.
- Cukier, R.I, H.B Levine and K.E Shuler (1978). 'Nonlinear sensitivity analysis of multiparameter model systems'. In: *Journal of Computational Physics* 26.1, pp. 1–42. ISSN: 00219991. DOI: [10.1016/0021-9991\(78\)90097-9](https://doi.org/10.1016/0021-9991(78)90097-9). URL: <http://linkinghub.elsevier.com/retrieve/pii/0021999178900979>.
- Das, Anjana, Domenico Rossetti di Valdalbero and Maria R. Virdis (2007). 'ACROPOLIS: An example of international collaboration in the field of energy modelling to support greenhouse gases mitigation policies'. In: *Energy Policy* 35.2, pp. 763–771. ISSN: 03014215. DOI: [10.1016/j.enpol.2006.03.004](https://doi.org/10.1016/j.enpol.2006.03.004). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421506001133>.
- Davies, Gareth et al. (2014). 'UKERC Energy Strategy Under Uncertainties Identifying techniques for managing uncertainty in the energy sector Working Paper'.
- DeCarolis, Joseph F, Kevin Hunter and Sarat Sreepathi (2012). 'Multi-stage stochastic optimization of a simple energy system'. In: *International Energy Workshop*, pp. 1–13.
- Department of Energy and Climate Change (2011). *Bioenergy Strategy Analytical Annex*. Tech. rep. April, pp. 1–96. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/48338/5136-bioenergy-strategy-analytical-annex.pdf<https://www.gov.uk/government/publications/uk-bioenergy-strategy>.
- Dixit, Avinash K and Robert S Pindyck (1994). *Investment under Uncertainty*. Princeton University Press. ISBN: 9780691034102. URL: http://books.google.co.uk/books?id=VahsELa{_}qC8C.
- Dow, S. C. (2004). 'Uncertainty and monetary policy'. In: *Oxford Economic Papers* 56.3, pp. 539–561. ISSN: 1464-3812. DOI: [10.1093/oep/gpf052](https://doi.org/10.1093/oep/gpf052). URL: <http://oep.oupjournals.org/cgi/doi/10.1093/oep/gpf052>.
- Dupacova, Jitka (2006). 'Optimization under exogenous and endogenous uncertainty'. URL: <http://scholar.google.com/scholar?hl=en{\&}btnG=Search{\&}q=intitle:Optimization+under+Exogenous+and+Endogenous+Uncertainty{\#}0>.
- Eckhause, Jeremy M., Danny R. Hughes and Steven a. Gabriel (2009). 'Evaluating real options for mitigating technical risk in public sector R&D acquisitions'. In: *International Journal of Project Management* 27.4, pp. 365–377. ISSN: 02637863. DOI: [10.1016/j.ijproman.2008.05.015](https://doi.org/10.1016/j.ijproman.2008.05.015). URL: <http://linkinghub.elsevier.com/retrieve/pii/S026378630800080X>.
- Eckhause, Jeremy and Johannes Herold (2014). 'Using real options to determine optimal funding strategies for CO₂ capture, transport

- and storage projects in the European Union'. In: *Energy Policy* 66, pp. 115–134. ISSN: 03014215. DOI: [10.1016/j.enpol.2013.11.030](https://doi.org/10.1016/j.enpol.2013.11.030). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421513011373>.
- Edmonds, J. and J. Reilly (1983). *A long-term global energy-economic model of carbon dioxide release from fossil fuel use*. DOI: [10.1016/0140-9883\(83\)90014-2](https://doi.org/10.1016/0140-9883(83)90014-2).
- Endo, E (2007). 'Market penetration analysis of fuel cell vehicles in Japan by using the energy system model MARKAL'. In: *International Journal of Hydrogen Energy* 32.10-11, pp. 1347–1354. ISSN: 03603199. DOI: [10.1016/j.ijhydene.2006.10.015](https://doi.org/10.1016/j.ijhydene.2006.10.015). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360319906004915>.
- Endo, E and M Ichinohe (2006). 'Analysis on market deployment of photovoltaics in Japan by using energy system model MARKAL'. In: *Solar Energy Materials and Solar Cells* 90.18-19, pp. 3061–3067. ISSN: 09270248. DOI: [10.1016/j.solmat.2006.06.049](https://doi.org/10.1016/j.solmat.2006.06.049). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0927024806002182>.
- Fawcett, Allen A et al. (2009). 'Overview of EMF 22 U.S. transition scenarios'. In: *Energy Economics* 31, S198–S211. ISSN: 01409883. DOI: [10.1016/j.eneco.2009.10.015](https://doi.org/10.1016/j.eneco.2009.10.015). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0140988309001984>.
- Felli, JC and GB Hazen (1999). 'A Bayesian approach to sensitivity analysis'. In: *Health Economics* 268.December 1998, pp. 263–268. DOI: [10.1002/\(SICI\)1099-1050\(199905\)8:3<263::AID-HEC426>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-1050(199905)8:3<263::AID-HEC426>3.0.CO;2-S). URL: [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1099-1050\(199905\)8:3{\\%}3C263::AID-HEC426{\\%}3E3.0.CO;2-S/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1099-1050(199905)8:3{\\%}3C263::AID-HEC426{\\%}3E3.0.CO;2-S/abstract).
- Fishbone, Leslie G. and Harold Abilock (1981). 'Markal, a linear-programming model for energy systems analysis: Technical description of the bnl version'. In: *International Journal of Energy Research* 5.4, pp. 353–375. ISSN: 0363907X. DOI: [10.1002/er.4440050406](https://doi.org/10.1002/er.4440050406). URL: <http://doi.wiley.com/10.1002/er.4440050406>.
- Flechsigt, M. et al. (2012). 'The Multi-Run Simulation Environment SimEnv: User Guide for Version 3.01'. Telegrafenberg, Potsdam, Germany.
- Goel, Vikas and Ignacio E Grossmann (2004). 'A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves'. In: 28, pp. 1409–1429. DOI: [10.1016/j.compchemeng.2003.10.005](https://doi.org/10.1016/j.compchemeng.2003.10.005).
- Goel, Vikas and Ignacio E Grossmann (2006). 'A Class of stochastic programs with decision dependent uncertainty'. In: *Mathematical Programming* 108.2-3, pp. 355–394. ISSN: 0025-5610. DOI: [10.1007/s10107-006-0715-7](https://doi.org/10.1007/s10107-006-0715-7). URL: <http://link.springer.com/10.1007/s10107-006-0715-7>.

- Golodnikov, A., A. Gritsevskii and S. Messner (1995). 'A stochastic version of the dynamic linear programming model MESSAGE III'. Laxenburg, Austria.
- Gül, Timur et al. (2009). 'An energy-economic scenario analysis of alternative fuels for personal transport using the Global Multi-regional MARKAL model (GMM)'. In: *Energy* 34.10, pp. 1423–1437. ISSN: 03605442. DOI: [10.1016/j.energy.2009.04.010](https://doi.org/10.1016/j.energy.2009.04.010). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360544209001078>.
- Gupta, Vijay and IE Grossmann (2011a). 'Offshore Oilfield Development Planning under Uncertainty and Fiscal Considerations'. In: *Optimization and Analytics in the Oil ...* Pp. 1–43. DOI: [10.1.1.214.8699](https://doi.org/10.1.1.214.8699). URL: http://egon.cheme.cmu.edu/Papers/GuptaGrossmannOptimization{_}and{_}Analytics{_}in{_}the{_}Oil{_}and{_}GasIndustry{_}v5.pdf.
- Gupta, Vijay and Ignacio E. Grossmann (2011b). 'Solution strategies for multistage stochastic programming with endogenous uncertainties'. In: *Computers & Chemical Engineering* 35.11, pp. 2235–2247. ISSN: 00981354. DOI: [10.1016/j.compchemeng.2010.11.013](https://doi.org/10.1016/j.compchemeng.2010.11.013). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0098135410003601>.
- Gupta, Vijay and Ignacio E Grossmann (2014). 'A new decomposition algorithm for multistage stochastic programs with endogenous uncertainties'. In: *Computers and Chemical Engineering* 62.1997, pp. 62–79. ISSN: 00981354. DOI: [10.1016/j.compchemeng.2013.11.011](https://doi.org/10.1016/j.compchemeng.2013.11.011).
- HM Government (2008). *Climate Change Act 2008*. URL: http://www.legislation.gov.uk/ukpga/2008/27/pdfs/ukpga{_}20080027{_}en.pdf.
- Hamby, D M (1994). 'A review of techniques for parameter sensitivity analysis of environmental models'. In: *Environmental Monitoring and Assessment* 32.2, pp. 135–154. ISSN: 0167-6369. DOI: [10.1007/BF00547132](https://doi.org/10.1007/BF00547132). URL: <http://deepblue.lib.umich.edu/handle/2027.42/42691http://link.springer.com/10.1007/BF00547132>.
- Hannart, Alexis et al. (2013). 'Disconcerting learning on climate sensitivity and the uncertain future of uncertainty'. In: *Climatic Change*, pp. 585–601. ISSN: 0165-0009. DOI: [10.1007/s10584-013-0770-z](https://doi.org/10.1007/s10584-013-0770-z). URL: <http://link.springer.com/10.1007/s10584-013-0770-z>.
- Hartley, K. (1965). 'The Learning Curve and Its Application to the Aircraft Industry'. In: *The Journal of Industrial Economics* 13.2, p. 122. ISSN: 00221821. DOI: [10.2307/2097667](https://doi.org/10.2307/2097667). URL: <http://www.jstor.org/stable/2097667?origin=crossref>.
- Heaton, Christopher and Rod Davies (2010). 'Modelling Low-Carbon UK Energy System Design through 2050 in a Collaboration of Industry and the Public Sector'. In: *IPMU*. 1, pp. 709–718.
- Held, Hermann et al. (2009). 'Efficient climate policies under technology and climate uncertainty'. In: *Energy Economics* 31, S50–S61.

- ISSN: 01409883. DOI: [10.1016/j.eneco.2008.12.012](https://doi.org/10.1016/j.eneco.2008.12.012). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0140988308001928>.
- Herman, J. D. et al. (2013). 'From maps to movies: High-resolution time-varying sensitivity analysis for spatially distributed watershed models'. In: *Hydrology and Earth System Sciences* 17.12, pp. 5109–5125. ISSN: 10275606. DOI: [10.5194/hess-17-5109-2013](https://doi.org/10.5194/hess-17-5109-2013).
- Herman, Jon et al. (2015). *SALib: Sensitivity Analysis Library in Python (Numpy)*. Contains Sobol, Morris, and FAST methods. DOI: [10.5281/zenodo.31316](https://doi.org/10.5281/zenodo.31316). URL: <http://salib.github.io/SALib/>.
- Hu, Ming-Che and Benjamin F. Hobbs (2010b). 'Analysis of multi-pollutant policies for the U.S. power sector under technology and policy uncertainty using MARKAL'. In: *Energy*. ISSN: 03605442. DOI: [10.1016/j.energy.2010.07.001](https://doi.org/10.1016/j.energy.2010.07.001). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360544210003658>.
- Hu, Ming-Che and Benjamin F. Hobbs (2010a). 'Analysis of multi-pollutant policies for the U.S. power sector under technology and policy uncertainty using MARKAL'. In: *Energy* 35.12, pp. 5430–5442. ISSN: 03605442. DOI: [10.1016/j.energy.2010.07.001](https://doi.org/10.1016/j.energy.2010.07.001). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360544210003658>.
- Hughes, Nick, Neil Strachan and Robert Gross (2012). 'The structure of uncertainty in future low carbon pathways'. In: *Energy Policy*, pp. 1–10. ISSN: 03014215. DOI: [10.1016/j.enpol.2012.04.028](https://doi.org/10.1016/j.enpol.2012.04.028). URL: <http://linkinghub.elsevier.com/retrieve/pii/S030142151200328X>.
- Huss, William R (1988). 'A move toward scenario analysis'. In: *International Journal of Forecasting* 4.3, pp. 377–388. ISSN: 01692070. DOI: [10.1016/0169-2070\(88\)90105-7](https://doi.org/10.1016/0169-2070(88)90105-7). URL: <http://linkinghub.elsevier.com/retrieve/pii/0169207088901057>.
- IPCC (2000). 'Quantifying Uncertainties in Practice'. In: *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. Ed. by J. Penman and S. Habetsion. IPCC. Chap. 6. ISBN: 9784887880009. URL: http://www.ipcc-nggip.iges.or.jp/public/gp/english/6{_}Uncertainty.pdf.
- Ichinohe, M and E Endo (2006). 'Analysis of the vehicle mix in the passenger-car sector in Japan for CO₂ emissions reduction by a MARKAL model'. In: *Applied Energy* 83.10, pp. 1047–1061. ISSN: 03062619. DOI: [10.1016/j.apenergy.2005.08.002](https://doi.org/10.1016/j.apenergy.2005.08.002). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0306261905001091>.
- International Energy Agency (2012). *Energy Technology Perspectives 2012*. Energy Technology Perspectives. OECD Publishing, p. 690. ISBN: 9789264174887. DOI: [10.1787/energy_tech-2012-en](https://doi.org/10.1787/energy_tech-2012-en). URL: http://www.oecd-ilibrary.org/energy/energy-technology-perspectives-2012{_}energy{_}tech-2012-en.
- Ireland, Gareth et al. (2015). 'Addressing the ability of a land biosphere model to predict key biophysical vegetation characterisation parameters with Global Sensitivity Analysis'. In: *Environ-*

- mental Modelling & Software* 65, pp. 94–107. ISSN: 13648152. DOI: [10.1016/j.envsoft.2014.11.010](https://doi.org/10.1016/j.envsoft.2014.11.010). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364815214003363>.
- Jebaraj, S and S Iniyan (2006). 'A review of energy models'. In: *Renewable and Sustainable Energy Reviews* 10.4, pp. 281–311. ISSN: 13640321. DOI: [10.1016/j.rser.2004.09.004](https://doi.org/10.1016/j.rser.2004.09.004). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364032104001261>.
- Jiang, B et al. (2008). 'The future of natural gas consumption in Beijing, Guangdong and Shanghai: An assessment utilizing MARKAL'. In: *Energy Policy* 36.9, pp. 3286–3299. ISSN: 03014215. DOI: [10.1016/j.enpol.2008.04.031](https://doi.org/10.1016/j.enpol.2008.04.031). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421508002188>.
- Jonsbraten, Tore Wiig (1998). 'Optimization Models for Petroleum Field Exploitation'. PhD thesis. Norwegian School of Economics and Business Administration. URL: <http://hdl.handle.net/11250/162551>.
- Jonsbraten, Tore Wiig, Roger J-b Wets and David L Woodruff (1998). 'A class of stochastic programs with decision dependent random elements'. In: 82, pp. 83–106.
- Kahn, Herman (1960). *On Thermonuclear War*. 2nd. New Jersey: Princeton, p. 651.
- Kann, Antje and John P. Weyant (2000). 'Approaches for performing uncertainty analysis in large-scale energy/economic policy models'. In: *Environmental Modeling and Assessment* 5.1, pp. 29–46. ISSN: 1420-2026. URL: <http://www.springerlink.com/index/N96V547567446488.pdf>.
- Kannan, Ramachandran (2009). 'Uncertainties in key low carbon power generation technologies – Implication for UK decarbonisation targets'. In: *Applied Energy* 86.10, pp. 1873–1886. ISSN: 03062619. DOI: [10.1016/j.apenergy.2009.02.014](https://doi.org/10.1016/j.apenergy.2009.02.014). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0306261909000579>.
- Kannan, Ramachandran and Neil Strachan (2009). 'Modelling the UK residential energy sector under long-term decarbonisation scenarios: Comparison between energy systems and sectoral modelling approaches'. In: *Applied Energy* 86.4, pp. 416–428. ISSN: 03062619. DOI: [10.1016/j.apenergy.2008.08.005](https://doi.org/10.1016/j.apenergy.2008.08.005). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0306261908001992>.
- Kanudia, Amit (1998). 'Modelling of Uncertainties and Price Elastic Demands in Energy-environment Planning for India'. In: *Omega* 26.3, pp. 409–423. ISSN: 03050483. DOI: [10.1016/S0305-0483\(97\)00071-6](https://doi.org/10.1016/S0305-0483(97)00071-6). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0305048397000716>.
- Kanudia, Amit and Richard Loulou (1999). 'Advanced bottom-up modelling for national and regional energy planning in response to climate change'. In: *International Journal of Environment and Pol-*

- lution 12.2/3, pp. 191–216. URL: <http://inderscience.metapress.com/index/q9bfvq5y63qlx9pd.pdf>.
- Kaplan, S (2000). ‘Combining probability distributions from experts in risk analysis’. In: *Risk analysis : an official publication of the Society for Risk Analysis* 20.2, pp. 155–6. ISSN: 1539-6924. URL: <http://www.ncbi.nlm.nih.gov/pubmed/10859775>.
- Keith, David W. (1996). ‘When is it appropriate to combine expert judgments?’ In: *Climatic Change* 33.2, pp. 139–143. ISSN: 0165-0009. DOI: 10.1007/BF00140244. URL: <http://www.springerlink.com/index/10.1007/BF00140244>.
- Keller, Klaus, Benjamin M Bolker and David F Bradford (2004). ‘Uncertain climate thresholds and optimal economic growth’. In: *Journal of Environmental Economics and Management* 48.1, pp. 723–741. ISSN: 00950696. DOI: 10.1016/j.jeem.2003.10.003. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0095069603001244>.
- Keppo, Ilkka and Bob Zwaan (2011). ‘The Impact of Uncertainty in Climate Targets and CO₂ Storage Availability on Long-Term Emissions Abatement’. In: *Environmental Modeling & Assessment* 17.1-2, pp. 177–191. ISSN: 1420-2026. DOI: 10.1007/s10666-011-9283-1. URL: <http://www.springerlink.com/index/10.1007/s10666-011-9283-1>.
- Kesicki, Fabian (2012). ‘Marginal Abatement Cost Curves: Combining Energy System Modelling and Decomposition Analysis’. In: *Environmental Modeling & Assessment* 18.1, pp. 27–37. ISSN: 1420-2026. DOI: 10.1007/s10666-012-9330-6. URL: <http://www.springerlink.com/index/10.1007/s10666-012-9330-6>.
- Kesicki, Fabian and Neil Strachan (2011a). ‘Marginal abatement cost (MAC) curves: confronting theory and practice’. In: *Environmental Science & Policy* 14.8, pp. 1195–1204. ISSN: 14629011. DOI: 10.1016/j.envsci.2011.08.004. URL: <http://linkinghub.elsevier.com/retrieve/pii/S1462901111001377>.
- Kesicki, Fabian and G Anandarajah (2011b). ‘The role of energy-service demand reduction in global climate change mitigation: Combining energy modelling and decomposition analysis’. In: *Energy Policy*. ISSN: 03014215. DOI: 10.1016/j.enpol.2011.08.043. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421511006409>.
- Kesicki, Fabian and Paul Ekins (2012). ‘Marginal abatement cost curves: a call for caution’. In: *Climate Policy* 12.2, pp. 219–236. ISSN: 1469-3062. DOI: 10.1080/14693062.2011.582347. URL: <http://www.tandfonline.com/doi/abs/10.1080/14693062.2011.582347>.
- Knight, FH (1921). *Risk, uncertainty and profit*. URL: http://books.google.com/books?hl=en&lr={&}id=Ntom6{_}pFQMCC{\&}oi=fnd{\&}pg=PR7{\&}dq=Risk,+Uncertainty,+and+Profit{\&}ots=5FJG{_}uw1e2{\&}sig={_}0y8LzWmpCVkgcgt3cP8400V5{_}A.

- Kolstad, Charles D. (1996). 'Learning and Stock Effects in Environmental Regulation: The Case of Greenhouse Gas Emissions'. In: *Journal of Environmental Economics and Management* 31.1, pp. 1–18. ISSN: 00950696. DOI: [10.1006/jeeem.1996.0028](https://doi.org/10.1006/jeeem.1996.0028). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0095069696900285>.
- Kramer, Gert Jan and Martin Haigh (2009). 'No quick switch to low-carbon energy.' In: *Nature* 462.7273, pp. 568–9. ISSN: 1476-4687. DOI: [10.1038/462568a](https://doi.org/10.1038/462568a). URL: <http://www.ncbi.nlm.nih.gov/pubmed/19956240>.
- Kriegler, Elmar (2009). 'Updating under unknown unknowns: An extension of Bayes' rule'. In: *International Journal of Approximate Reasoning* 50.4, pp. 583–596. ISSN: 0888613X. DOI: [10.1016/j.ijar.2008.09.005](https://doi.org/10.1016/j.ijar.2008.09.005). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0888613X08001631>.
- Kriegler, Elmar et al. (2015). 'Diagnostic indicators for integrated assessment models of climate policy'. In: *Technological Forecasting and Social Change* 90, pp. 45–61. ISSN: 00401625. DOI: [10.1016/j.techfore.2013.09.020](https://doi.org/10.1016/j.techfore.2013.09.020). URL: <http://dx.doi.org/10.1016/j.techfore.2013.09.020><http://linkinghub.elsevier.com/retrieve/pii/S0040162513002576>.
- Krukanont, Pongsak and Tetsuo Tezuka (2007). 'Implications of capacity expansion under uncertainty and value of information: The near-term energy planning of Japan'. In: *Energy* 32.10, pp. 1809–1824. ISSN: 03605442. DOI: [10.1016/j.energy.2007.02.003](https://doi.org/10.1016/j.energy.2007.02.003). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360544207000333>.
- Krzyzanowski, Daniel a., Socrates Kypreos and Leonardo Barreto (2007). 'Supporting hydrogen based transportation: case studies with Global MARKAL Model'. In: *Computational Management Science* 5.3, pp. 207–231. ISSN: 1619-697X. DOI: [10.1007/s10287-007-0040-5](https://doi.org/10.1007/s10287-007-0040-5). URL: <http://www.springerlink.com/index/10.1007/s10287-007-0040-5>.
- Kucherenko, S., S. Tarantola and P. Annoni (2012). 'Estimation of global sensitivity indices for models with dependent variables'. In: *Computer Physics Communications* 183.4, pp. 937–946. ISSN: 00104655. DOI: [10.1016/j.cpc.2011.12.020](https://doi.org/10.1016/j.cpc.2011.12.020). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0010465511004085>.
- Labriet, Maryse, Richard Loulou and Amit Kanudia (2008). 'Is a 2 degrees Celsius warming achievable under high uncertainty? Analysis with the TIMES integrated assessment model'. URL: <http://www.etsap.org/TIAM/f/1/stochastic/climatepolicy.pdf>.
- Labriet, Maryse, Richard Loulou and Amit Kanudia (2010). 'Modeling Uncertainty in a Large Scale Integrated Energy-Climates Model'. In: *Uncertainty and Environmental Decision Making*. Ed. by Jerzy A. Filar and Alain Haurie. Vol. 138. International Series in Operations Research & Management Science. Boston, MA: Springer

- US, pp. 51–77. ISBN: 978-1-4419-1128-5. DOI: [10.1007/978-1-4419-1129-2](https://doi.org/10.1007/978-1-4419-1129-2). URL: <http://www.springerlink.com/content/h5124t576t186480/http://www.gerad.ca/fichiers/cahiers/G-2009-06.pdf>.
- Lagerwall, Gareth et al. (2014). 'Global uncertainty and sensitivity analysis of a spatially distributed ecological model'. In: *Ecological Modelling* 275, pp. 22–30. ISSN: 03043800. DOI: [10.1016/j.ecolmodel.2013.12.010](https://doi.org/10.1016/j.ecolmodel.2013.12.010).
- Lange, Andreas and Nicolas Treich (2008). 'Uncertainty, learning and ambiguity in economic models on climate policy: some classical results and new directions'. In: *Climatic Change* 89.1-2, pp. 7–21. ISSN: 0165-0009. DOI: [10.1007/s10584-008-9401-5](https://doi.org/10.1007/s10584-008-9401-5). URL: <http://www.springerlink.com/index/10.1007/s10584-008-9401-5>.
- Larsson, Tomas (1993). 'Developing strategies for robust energy systems. II: Application to CO₂ risk management'. In: *International Journal of Energy Research* 17.6, pp. 515–535. ISSN: 0363907X. DOI: [10.1002/er.4440170607](https://doi.org/10.1002/er.4440170607). URL: <http://doi.wiley.com/10.1002/er.4440170607>.
- Larsson, Tomas and Clas-Otto Wene (1993). 'Developing strategies for robust energy systems. I: Methodology'. In: *International Journal of Energy Research* 17.6, pp. 503–513. ISSN: 0363907X. DOI: [10.1002/er.4440170606](https://doi.org/10.1002/er.4440170606). URL: <http://doi.wiley.com/10.1002/er.4440170606>.
- Lempert, Robert J (2002). 'A new decision sciences for complex systems.' In: *Proceedings of the National Academy of Sciences of the United States of America* 99.10/3, pp. 7309–7313. ISSN: 0027-8424. DOI: [10.1073/pnas.082081699](https://doi.org/10.1073/pnas.082081699). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=128602&tool=pmcentrez&rendertype=abstract>.
- Lempert, Robert J, Steven W Popper and Steven C Bankes (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*. Santa Monica, California: RAND Corporation. ISBN: 0833032755.
- Lempert, Robert J and Myles T Collins (2007). 'Managing the risk of uncertain threshold responses: comparison of robust, optimum, and precautionary approaches.' In: *Risk analysis : an official publication of the Society for Risk Analysis* 27.4, pp. 1009–26. ISSN: 0272-4332. DOI: [10.1111/j.1539-6924.2007.00940.x](https://doi.org/10.1111/j.1539-6924.2007.00940.x). URL: <http://www.ncbi.nlm.nih.gov/pubmed/17958508>.
- Lorenz, Alexander et al. (2011). 'Anticipating Climate Threshold Damages'. In: *Environmental Modeling and Assessment* 17.1-2, pp. 163–175. ISSN: 1420-2026. DOI: [10.1007/s10666-011-9282-2](https://doi.org/10.1007/s10666-011-9282-2). URL: <http://www.springerlink.com/index/10.1007/s10666-011-9282-2>.
- Loulou, Richard and Amit Kanudia (1999). 'Minimax regret strategies for greenhouse gas abatement: methodology and application'. In:

- Operations Research Letters* 25.5, pp. 219–230. ISSN: 01676377. DOI: [10.1016/S0167-6377\(99\)00049-8](https://doi.org/10.1016/S0167-6377(99)00049-8). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0167637799000498>.
- Loulou, Richard, Maryse Labriet and Amit Kanudia (2009). ‘Deterministic and stochastic analysis of alternative climate targets under differentiated cooperation regimes’. In: *Energy Economics* 31, S131–S143. ISSN: 01409883. DOI: [10.1016/j.eneco.2009.06.012](https://doi.org/10.1016/j.eneco.2009.06.012). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0140988309001108>.
- Low Carbon Innovation Coordination Group (2014). *Coordinating Low Carbon Technology Innovation Support The LCICG’s Strategic Framework*. Tech. rep. February. Low Carbon Innovation Coordination Group, p. 83.
- McJeon, Haewon C. et al. (2011). ‘Technology interactions among low-carbon energy technologies: What can we learn from a large number of scenarios?’ In: *Energy Economics* 33.4, pp. 619–631. ISSN: 01409883. DOI: [10.1016/j.eneco.2010.10.007](https://doi.org/10.1016/j.eneco.2010.10.007). URL: <http://dx.doi.org/10.1016/j.eneco.2010.10.007>.
- McManus, Hugh and Daniel Hastings (2006). ‘A framework for understanding uncertainty and its mitigation and exploitation in complex systems’. In: *IEEE Engineering Management Review* 34.3, pp. 81–81. ISSN: 0360-8581. DOI: [10.1109/EMR.2006.261384](https://doi.org/10.1109/EMR.2006.261384). URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4032565>.
- Meinshausen, Malte et al. (2011). ‘The RCP greenhouse gas concentrations and their extensions from 1765 to 2300’. In: *Climatic Change* 109.1-2, pp. 213–241. ISSN: 0165-0009. DOI: [10.1007/s10584-011-0156-z](https://doi.org/10.1007/s10584-011-0156-z). URL: <http://www.springerlink.com/index/10.1007/s10584-011-0156-z>.
- Mercier, Luc and Pascal Van Hentenryck (2008). ‘Amsaa: A Multistep Anticipatory Algorithm for Online Stochastic Combinatorial Optimization’. In: ... *for Combinatorial Optimization Problems*, pp. 173–187. DOI: [10.1007/978-3-540-68155-7_15](https://doi.org/10.1007/978-3-540-68155-7_15). URL: http://link.springer.com/chapter/10.1007/978-3-540-68155-7_15http://link.springer.com/10.1007/978-3-540-68155-7_15.
- Messner, S., A. Golodnikov and A. Gritsevskii (1996). ‘A stochastic version of the dynamic linear programming model MESSAGE III’. In: *Energy* 21.9, pp. 775–784. ISSN: 03605442. DOI: [10.1016/0360-5442\(96\)00025-4](https://doi.org/10.1016/0360-5442(96)00025-4). URL: <http://linkinghub.elsevier.com/retrieve/pii/0360544296000254>.
- Miller, K (2003). ‘Scenarios, Real Options and Integrated Risk Management’. In: *Long Range Planning* 36.1, pp. 93–107. ISSN: 00246301. DOI: [10.1016/S0024-6301\(02\)00205-4](https://doi.org/10.1016/S0024-6301(02)00205-4). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0024630102002054>.

- Morgan, M. Granger, M. Henrion and M. Small (1992). *Uncertainty: A guide to dealing with uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, p. 332. ISBN: 9780521427449. URL: www.cambridge.org/9780521427449.
- Morris, Max D (1991). 'Factorial Sampling Plans for Preliminary Computational Experiments'. In: *Technometrics* 33, pp. 161–174. ISSN: 0040-1706. DOI: [10.2307/1269043](https://doi.org/10.2307/1269043). URL: <http://www.jstor.org/stable/1269043>.
- Morris, P.A. (1977). 'Combining expert judgments: A Bayesian approach'. In: *Management Science* 23.7, pp. 679–693. URL: <http://www.jstor.org/stable/2630848>.
- Nemet, Gregory F., Laura Diaz Anadon and Elena Verdolini (2016). 'Quantifying the Effects of Expert Selection and Elicitation Design on Experts' Confidence in Their Judgments About Future Energy Technologies'. In: *Risk Analysis* 4. ISSN: 15396924. DOI: [10.1111/risa.12604](https://doi.org/10.1111/risa.12604).
- Nordhaus, William D. (2014). 'The Perils of the Learning Model for Modeling Endogenous Technological Change'. In: *The Energy Journal* 35.1, pp. 1–14. ISSN: 01956574. DOI: [10.5547/01956574.35.1.1](https://doi.org/10.5547/01956574.35.1.1). URL: <http://www.iaee.org/en/publications/ejarticle.aspx?id=2546>.
- Nossent, Jiri, Pieter Elsen and Willy Bauwens (2011). 'Sobol' sensitivity analysis of a complex environmental model'. In: *Environmental Modelling & Software* 26.12, pp. 1515–1525. ISSN: 13648152. DOI: [10.1016/j.envsoft.2011.08.010](https://doi.org/10.1016/j.envsoft.2011.08.010). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364815211001939>.
- Nurminski, Evgeni and Todor Balabanov (1982). 'Decomposition of a large-scale energy model'. Laxenburg, Austria.
- O'Hagan, Anthony (1998). 'Eliciting expert beliefs in substantial practical applications [Read before The Royal Statistical Society at a meeting on 'Elicitation' on Wednesday, April 16th, 1997, the President, Professor A. F. M. Smith in the Chair]'. In: *Journal of the Royal Statistical Society: Series D (The Statistician)* 47.1, pp. 21–35. ISSN: 0039-0526. DOI: [10.1111/1467-9884.00114](https://doi.org/10.1111/1467-9884.00114). URL: <http://doi.wiley.com/10.1111/1467-9884.00114>.
- O'Hagan, Anthony (2006a). *Uncertain judgements: eliciting experts' probabilities*. London: John Wiley & Sons. ISBN: 0470029994.
- O'Hagan, Anthony (2006b). *Uncertain judgements: eliciting experts' probabilities*. London: John Wiley & Sons. ISBN: 0470029994.
- O'Hagan, Anthony and Jeremy E. Oakley (2004). 'Probability is perfect, but we can't elicit it perfectly'. In: *Reliability Engineering & System Safety* 85.1-3, pp. 239–248. ISSN: 09518320. DOI: [10.1016/j.res.2004.03.014](https://doi.org/10.1016/j.res.2004.03.014). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0951832004000638>.
- Oakley, Jeremy E. and Anthony O'Hagan (2004). 'Probabilistic sensitivity analysis of complex models: a Bayesian approach'. In: *Journal*

- of the Royal Statistical Society: Series B (Statistical Methodology) 66.3, pp. 751–769. ISSN: 1369-7412. DOI: [10.1111/j.1467-9868.2004.05304.x](https://doi.org/10.1111/j.1467-9868.2004.05304.x). URL: <http://doi.wiley.com/10.1111/j.1467-9868.2004.05304.x>.
- Oppenheimer, Michael, Brian C. O'Neill and Mort Webster (2008). 'Negative learning'. In: *Climatic Change* 89.1-2, pp. 155–172. ISSN: 0165-0009. DOI: [10.1007/s10584-008-9405-1](https://doi.org/10.1007/s10584-008-9405-1). URL: <http://www.springerlink.com/index/10.1007/s10584-008-9405-1>.
- Painter, Brett D.M. and Mark W. Milke (2007). 'Comparison of Factorial and Scenario Analysis Methods for Assessing Uncertainty in the Design of Permeable Reactive Barriers'. In: *Ground Water Monitoring & Remediation* 27.3, pp. 102–110. ISSN: 1069-3629. DOI: [10.1111/j.1745-6592.2007.00152.x](https://doi.org/10.1111/j.1745-6592.2007.00152.x). URL: <http://doi.wiley.com/10.1111/j.1745-6592.2007.00152.x>.
- Plischke, Elmar, Emanuele Borgonovo and Curtis L. Smith (2013). 'Global sensitivity measures from given data'. In: *European Journal of Operational Research* 226.3, pp. 536–550. ISSN: 03772217. DOI: [10.1016/j.ejor.2012.11.047](https://doi.org/10.1016/j.ejor.2012.11.047). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0377221712008995>.
- Popp, David and William D. Nordhaus (1997). 'What is the Value of Scientific Knowledge? An Application to Global Warming Using the PRICE Model'. In: *The Energy Journal* 18.1, pp. 1–45. URL: <http://www.jstor.org/stable/41322716?origin=JSTOR-pdf>.
- Poropudas, Jirka, Jouni Pousi and Kai Virtanen (2011). 'Multiple input and multiple output simulation metamodeling using Bayesian networks'. In: *Proceedings of the 2011 Winter Simulation Conference (WSC)*. Ed. by S. Jain et al. IEEE, pp. 569–580. ISBN: 978-1-4577-2109-0. DOI: [10.1109/WSC.2011.6147786](https://doi.org/10.1109/WSC.2011.6147786). URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6147786>.
- Powell, WB (2011). *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley Series in Probability and Statistics. Wiley. ISBN: 9781118029169. URL: <http://books.google.co.uk/books?id=RS53Dyx8wcUC>.
- Powell, WB and I O Ryzhov (2012). *Optimal Learning*. Wiley Series in Probability and Statistics. John Wiley & Sons. ISBN: 9780470596692. URL: <http://books.google.co.uk/books?id=hnsVMbx5H0AC>.
- Pye, Steve, Will Usher and Neil Strachan (2014). 'The uncertain but critical role of demand reduction in meeting long-term energy decarbonisation targets'. In: *Energy Policy* 73.2000, pp. 575–586. ISSN: 03014215. DOI: [10.1016/j.enpol.2014.05.025](https://doi.org/10.1016/j.enpol.2014.05.025). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421514003061>.
- Pye, Steve, Nagore Sabio and Neil Strachan (2015). 'An integrated systematic analysis of uncertainties in UK energy transition pathways'. In: *Energy Policy*, pp. 1–12. ISSN: 0301-4215. DOI: [10.1016/j.enpol.2014.12.031](https://doi.org/10.1016/j.enpol.2014.12.031). URL: <http://dx.doi.org/10.1016/j.enpol.2014.12.031>.

- Rafaj, Peter and Socrates Kypreos (2007). 'Internalisation of external cost in the power generation sector: Analysis with Global Multi-regional MARKAL model'. In: *Energy Policy* 35.2, pp. 828–843. ISSN: 03014215. DOI: [10.1016/j.enpol.2006.03.003](https://doi.org/10.1016/j.enpol.2006.03.003). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421506001224>.
- Rath-Nagel, S (1982). 'Energy modelling for technology assessment: the MARKAL approach'. In: *Omega* 10.5, pp. 493–505. ISSN: 03050483. DOI: [10.1016/0305-0483\(82\)90006-8](https://doi.org/10.1016/0305-0483(82)90006-8). URL: <http://linkinghub.elsevier.com/retrieve/pii/0305048382900068>.
- Rhodes, Aidan, Jim Skea and Matthew Hannon (2014). 'The Global Surge in Energy Innovation'. In: *Energies* 7.9, pp. 5601–5623. ISSN: 1996-1073. DOI: [10.3390/en7095601](https://doi.org/10.3390/en7095601). URL: <http://www.mdpi.com/1996-1073/7/9/5601/>.
- Rosenberg, Eva et al. (2010). 'Market penetration analysis of hydrogen vehicles in Norwegian passenger transport towards 2050'. In: *International Journal of Hydrogen Energy* 35.14, pp. 7267–7279. ISSN: 03603199. DOI: [10.1016/j.ijhydene.2010.04.153](https://doi.org/10.1016/j.ijhydene.2010.04.153). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360319910008554>.
- Rotmans, Jan and M.B.A. van Asselt (2001a). 'Uncertainty in integrated assessment modelling: A labyrinthic path'. In: *Integrated Assessment* 2.2, pp. 43–55. URL: <http://www.springerlink.com/index/PUU57104W7412245.pdf>.
- Rotmans, Jan and M.B.A. van Asselt (2001b). 'Uncertainty management in integrated assessment modeling: towards a pluralistic approach.' In: *Environmental monitoring and assessment* 69.2, pp. 101–130. ISSN: 0167-6369. URL: <http://www.ncbi.nlm.nih.gov/pubmed/11465663>.
- Ruano, M. V. et al. (2012). 'An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors'. In: *Environmental Modelling and Software* 37, pp. 103–109. ISSN: 13648152. DOI: [10.1016/j.envsoft.2012.03.008](https://doi.org/10.1016/j.envsoft.2012.03.008). URL: <http://dx.doi.org/10.1016/j.envsoft.2012.03.008>.
- Saltelli, Andrea (2002). 'Making best use of model evaluations to compute sensitivity indices'. In: *Computer Physics Communications* 145.2, pp. 280–297. ISSN: 00104655. DOI: [10.1016/S0010-4655\(02\)00280-1](https://doi.org/10.1016/S0010-4655(02)00280-1). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0010465502002801>.
- Saltelli, Andrea, T.H. Andres and T. Homma (1995). 'Sensitivity analysis of model output. Performance of the iterated fractional factorial design method'. In: *Computational Statistics & Data Analysis* 20.4, pp. 387–407. ISSN: 01679473. DOI: [10.1016/0167-9473\(95\)92843-M](https://doi.org/10.1016/0167-9473(95)92843-M).
- Saltelli, Andrea and Ricardo Bolado (1998). 'An alternative way to compute Fourier amplitude sensitivity test (FAST)'. In: *Computa-*

- tional Statistics & Data Analysis* 26.4, pp. 445–460. ISSN: 01679473. DOI: [10.1016/S0167-9473\(97\)00043-1](https://doi.org/10.1016/S0167-9473(97)00043-1).
- Saltelli, Andrea, K. Chan and E. M. Scott (2000). *Sensitivity Analysis*. Wiley, p. 426. ISBN: 0471998923.
- Saltelli, Andrea et al. (2005). 'Sensitivity analysis for chemical models.' In: *Chemical Reviews* 105.7, p. 2811. ISSN: 0009-2665. DOI: [10.1021/cr040659d](https://doi.org/10.1021/cr040659d). URL: [papers2://publication/uuid/C649A15C-DDC4-4705-A66D-C42B34D88729](https://pubs.acs.org/doi/10.1021/cr040659d).
- Saltelli, Andrea et al. (2008a). *Global Sensitivity Analysis: The Primer*. Wiley. ISBN: 9780470725177. URL: <http://books.google.co.uk/books?id=wAssmt2vumgC>.
- Saltelli, Andrea et al. (2008b). 'Variance-Based Methods'. In: *Global Sensitivity Analysis. The Primer*. Chap. 4, pp. 155–174. ISBN: 9780470059975. DOI: [10.1002/9780470725184.ch4](https://doi.org/10.1002/9780470725184.ch4). URL: <http://onlinelibrary.wiley.com/doi/10.1002/9780470725184.ch4/summary>.
- Saltelli, Andrea and Paola Annoni (2010a). 'How to avoid a perfunctory sensitivity analysis'. In: *Environmental Modelling & Software* 25.12, pp. 1508–1517. ISSN: 13648152. DOI: [10.1016/j.envsoft.2010.04.012](https://doi.org/10.1016/j.envsoft.2010.04.012). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364815210001180>.
- Saltelli, Andrea et al. (2010b). 'Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index'. In: *Computer Physics Communications* 181.2, pp. 259–270. ISSN: 00104655. DOI: [10.1016/j.cpc.2009.09.018](https://doi.org/10.1016/j.cpc.2009.09.018). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0010465509003087>.
- Salvia, Monica et al. (2002). 'Waste management system optimisation for Southern Italy with MARKAL model'. In: *Resources, Conservation and Recycling* 34.2, pp. 91–106. ISSN: 09213449. DOI: [10.1016/S0921-3449\(01\)00095-7](https://doi.org/10.1016/S0921-3449(01)00095-7). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0921344901000957>.
- Santen, Nidhi R and Laura Diaz (2016). 'Balancing solar PV deployment and RD & D : A comprehensive framework for managing innovation uncertainty in electricity technology investment planning'. In: *Renewable and Sustainable Energy Reviews* 60, pp. 560–569. ISSN: 1364-0321. DOI: [10.1016/j.rser.2015.12.272](https://doi.org/10.1016/j.rser.2015.12.272). URL: <http://dx.doi.org/10.1016/j.rser.2015.12.272>.
- Schrattenholzer, Leo (1981). *The energy supply model MESSAGE*. Tech. rep. Laxenburg, Austria: IIASA, p. 40.
- Schulz, T et al. (2007). 'Assessing wood-based synthetic natural gas technologies using the SWISS-MARKAL model'. In: *Energy* 32.10, pp. 1948–1959. ISSN: 03605442. DOI: [10.1016/j.energy.2007.03.006](https://doi.org/10.1016/j.energy.2007.03.006). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360544207000564>.
- Scott, Michael J. et al. (2014). 'Evaluating sub-national building-energy efficiency policy options under uncertainty: Efficient sensitivity testing of alternative climate, technological, and socioeconomic

- futures in a regional integrated-assessment model'. In: *Energy Economics* 43, pp. 22–33. ISSN: 01409883. DOI: [10.1016/j.eneco.2014.01.012](https://doi.org/10.1016/j.eneco.2014.01.012). URL: <http://dx.doi.org/10.1016/j.eneco.2014.01.012>.
- Seebregts, A J et al. (1999). *Endogenous technological change in energy system models*. Tech. rep. Laxenburg, Austria: IIASA.
- Shittu, Ekundayo and Erin Baker (2010). 'Optimal Energy R&D Portfolio Investments in Response to a Carbon Tax'. In: *IEEE Transactions on Engineering Management* 57.4, pp. 547–559. ISSN: 0018-9391. DOI: [10.1109/TEM.2009.2023107](https://doi.org/10.1109/TEM.2009.2023107). URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5208293>.
- Shlyakhter, Alexander I et al. (1994). 'Quantifying the credibility of energy projections from trends in past data'. In: *Energy Policy* 22.2, pp. 119–130. ISSN: 03014215. DOI: [10.1016/0301-4215\(94\)90129-5](https://doi.org/10.1016/0301-4215(94)90129-5). URL: <http://www.sciencedirect.com/science/article/pii/0301421594901295><http://linkinghub.elsevier.com/retrieve/pii/0301421594901295>.
- Skea, Jim (2014). 'The renaissance of energy innovation'. In: *Energy Environ. Sci.* 7.1, pp. 21–24. ISSN: 1754-5692. DOI: [10.1039/C3EE43034K](https://doi.org/10.1039/C3EE43034K). URL: <http://xlink.rsc.org/?DOI=C3EE43034K>.
- Sluijs, Jeroen P. van der et al. (2002). *Dutch National Research Programme on Global Air Uncertainty assessment of the IMAGE / TIMER B1 CO 2 emissions scenario , using the NUSAP method*. Tech. rep.
- Sluijs, Jeroen P van der et al. (2005). 'Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: the NUSAP system.' In: *Risk analysis : an official publication of the Society for Risk Analysis* 25.2, pp. 481–492. ISSN: 0272-4332. DOI: [10.1111/j.1539-6924.2005.00604.x](https://doi.org/10.1111/j.1539-6924.2005.00604.x).
- Sluijs, Jeroen Pieter van der (1997). 'Anchoring amid uncertainty: On the management of uncertainties in risk assessment of anthropogenic climate change'. PhD thesis, pp. 1–213.
- Smil, Vaclav (2000). 'Perils of Long-Range Energy Forecasting Reflections on Looking Far Ahead'. In: *Technological Forecasting and Social Change* 65.3, pp. 251–264. ISSN: 00401625. DOI: [10.1016/S0040-1625\(99\)00097-9](https://doi.org/10.1016/S0040-1625(99)00097-9). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0040162599000979>.
- Sobol', Il'ya Meerovich (1990). 'On sensitivity estimation for nonlinear mathematical models'. In: *Matematicheskoe Modelirovanie* 2.1, pp. 112–118. ISSN: 0234-0879.
- Socolow, Robert H. (2011). 'High-consequence outcomes and internal disagreements: Tell us more, please'. In: *Climatic Change* 108, pp. 775–790. ISSN: 01650009. DOI: [10.1007/s10584-011-0187-5](https://doi.org/10.1007/s10584-011-0187-5).
- Solak, Senay et al. (2010). 'Optimization of R&D project portfolios under endogenous uncertainty'. In: *European Journal of Operational Research* 207.1, pp. 420–433. ISSN: 03772217. DOI: [10.1016/j.ejor](https://doi.org/10.1016/j.ejor).

- 2010.04.032. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0377221710003516>.
- Spiegelhalter, David J and Hauke Riesch (2011a). 'Don't know, can't know: embracing deeper uncertainties when analysing risks.' In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 369.1956, pp. 4730–50. ISSN: 1364-503X. DOI: 10.1098/rsta.2011.0163. URL: <http://www.ncbi.nlm.nih.gov/pubmed/22042895>.
- Spiegelhalter, David J, M. Pearson and I. Short (2011b). 'Visualizing Uncertainty About the Future'. In: *Science* 333.6048, pp. 1393–1400. ISSN: 0036-8075. DOI: 10.1126/science.1191181. URL: <http://www.sciencemag.org/cgi/doi/10.1126/science.1191181>.
- Srivastava, Vibhava et al. (2014). 'Insights on geologic and vegetative controls over hydrologic behavior of a large complex basin – Global Sensitivity Analysis of an integrated parallel hydrologic model'. In: *Journal of Hydrology* 519, pp. 2238–2257. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.10.020. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0022169414008129>.
- Stirling, Andy (2007). 'Risk, precaution and science: towards a more constructive policy debate. Talking point on the precautionary principle.' In: *EMBO reports* 8.4, pp. 309–15. ISSN: 1469-221X. DOI: 10.1038/sj.embor.7400953. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1852772&tool=pmcentrez&rendertype=abstract>.
- Stirling, Andy (2010). 'Keep it complex'. In: *Nature* 468.7327, pp. 1029–1031. URL: <http://www.nature.com/nature/journal/v468/n7327/abs/4681029a.html>.
- Strachan, Neil (2010). 'Business-as-Unusual: Existing policies in energy model baselines'. In: *Energy Economics*. ISSN: 01409883. DOI: 10.1016/j.eneco.2010.10.009. URL: <http://linkinghub.elsevier.com/retrieve/pii/S014098831000191X>.
- Strachan, Neil and Ramachandran Kannan (2008a). 'Hybrid modelling of long-term carbon reduction scenarios for the UK'. In: *Energy Economics* 30.6, pp. 2947–2963. ISSN: 01409883. DOI: 10.1016/j.eneco.2008.04.009. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0140988308000613>.
- Strachan, Neil, T Foxon and J Fujino (2008b). *Modelling long-term scenarios for low carbon societies*. Climate Policy. Earthscan. ISBN: 9781844075942. URL: <http://books.google.co.uk/books?id=JzAMLQAACAAJ>.
- Strachan, Neil, Tim Foxon and Junichi Fujino (2008c). 'Policy implications from the Low-Carbon Society (LCS) modelling project'. In: *Climate Policy* 8.1, pp. 17–29. ISSN: 14693062. DOI: 10.3763/cpol.2007.0488. URL: <http://openurl.ingenta.com/content/xref?genre=article&issn=1469-3062&volume=8&issue=51&spage=17>.

- Strachan, Neil, Steve Pye and Ramachandran Kannan (2008d). 'The iterative contribution and relevance of modelling to UK energy policy'. In: *Energy Policy* 37, pp. 850–860. ISSN: 03014215. DOI: [10.1016/j.enpol.2008.09.096](https://doi.org/10.1016/j.enpol.2008.09.096). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421508005648>.
- Strachan, Neil, Steve Pye and Nick Hughes (2008e). 'The role of international drivers on UK scenarios of a low-carbon society'. In: *Climate Policy* 8.1, pp. 125–139. ISSN: 14693062. DOI: [10.3763/cpol.2007.0489](https://doi.org/10.3763/cpol.2007.0489). URL: <http://openurl.ingenta.com/content/xref?genre=article{\&}issn=1469-3062{\&}volume=8{\&}issue=51{\&}spage=125>.
- Strachan, Neil et al. (2009). 'Soft-linking energy systems and GIS models to investigate spatial hydrogen infrastructure development in a low-carbon UK energy system'. In: *International Journal of Hydrogen Energy* 34.2, pp. 642–657. ISSN: 03603199. DOI: [10.1016/j.ijhydene.2008.10.083](https://doi.org/10.1016/j.ijhydene.2008.10.083). URL: <http://linkinghub.elsevier.com/retrieve/pii/S036031990801402X>.
- Strachan, Neil and Will Usher (2011). 'Failure to achieve stringent carbon reduction targets in a second-best policy world'. In: *Climatic Change* 113.2, pp. 121–139. ISSN: 0165-0009. DOI: [10.1007/s10584-011-0267-6](https://doi.org/10.1007/s10584-011-0267-6). URL: <http://www.springerlink.com/index/10.1007/s10584-011-0267-6>.
- Tarhan, Bora, Ignacio E Grossmann and Vikas Goel (2009). 'Stochastic Programming Approach for the Planning of Offshore Oil or Gas Field Infrastructure under Decision-Dependent Uncertainty'. In: pp. 3078–3097.
- Tarhan, Bora, Ignacio E. Grossmann and Vikas Goel (2011). 'Computational strategies for non-convex multistage MINLP models with decision-dependent uncertainty and gradual uncertainty resolution'. In: *Annals of Operations Research* 203.1, pp. 141–166. ISSN: 0254-5330. DOI: [10.1007/s10479-011-0855-x](https://doi.org/10.1007/s10479-011-0855-x). URL: <http://link.springer.com/10.1007/s10479-011-0855-x>.
- Tetlock, Philip E (2005). *Expert Political Judgement*. Princeton University Press, p. 321. ISBN: 0691123020.
- Trutnevyte, Evelina et al. (2016). 'Energy scenario choices: Insights from a retrospective review of UK energy futures'. In: *Renewable and Sustainable Energy Reviews* 55, pp. 326–337. ISSN: 13640321. DOI: [10.1016/j.rser.2015.10.067](https://doi.org/10.1016/j.rser.2015.10.067). URL: <http://linkinghub.elsevier.com/retrieve/pii/S1364032115011466>.
- Tseng, Phillip, John Lee and Paul Friley (2005). 'A hydrogen economy: opportunities and challenges'. In: *Energy* 30.14, pp. 2703–2720. ISSN: 03605442. DOI: [10.1016/j.energy.2004.07.015](https://doi.org/10.1016/j.energy.2004.07.015). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0360544204003160>.
- Usher, Will and Neil Strachan (2010a). *UK MARKAL Modelling - Examining Decarbonisation Pathways in the 2020s on the Way to Meeting the 2050 Emissions Target*. Tech. rep. London, UK: Committee

- on Climate Change, pp. 1–159. URL: <http://downloads.theccc.org.uk/s3.amazonaws.com/4thBudget/CCCMARKALFinalReport-UCLNov10.pdf>.
- Usher, Will and Neil Strachan (2012). 'Critical mid-term uncertainties in long-term decarbonisation pathways'. In: *Energy Policy* 41, pp. 433–444. ISSN: 03014215. DOI: 10.1016/j.enpol.2011.11.004. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421511008810>.
- Usher, Will and Neil Strachan (2013). 'An expert elicitation of climate, energy and economic uncertainties'. In: *Energy Policy* 61, pp. 811–821. ISSN: 03014215. DOI: 10.1016/j.enpol.2013.06.110. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421513006174>.
- Usher, William and Neil Strachan (2010b). *UK MARKAL Modelling - Examining Decarbonisation Pathways in the 2020s on the Way to Meeting the 2050 Emissions Target*. Tech. rep., pp. 1–159. URL: <http://downloads.theccc.org.uk/s3.amazonaws.com/4thBudget/CCCMARKALFinalReport-UCLNov10.pdf>.
- Vayanos, Phebe and Daniel Kuhn (2011). 'Decision rules for information discovery in multi-stage stochastic programming'. In: pp. 7368–7373.
- Wack, Pierre (1985a). 'Scenarios: shooting the rapids'. In: *Harvard Business Review*.
- Wack, Pierre (1985b). 'Scenarios: uncharted waters ahead'. In: *Harvard Business Review*, pp. 73–89.
- Wagner, Harvey M. (1995). 'Global Sensitivity Analysis'. In: *Operations Research* 43.6, pp. 948–969. ISSN: 0030-364X. DOI: 10.1287/opre.43.6.948. URL: <http://pubsonline.informs.org/doi/abs/10.1287/opre.43.6.948>.
- Wainwright, Haruko M. et al. (2014). 'Making sense of global sensitivity analyses'. In: *Computers & Geosciences* 65, pp. 84–94. ISSN: 00983004. DOI: 10.1016/j.cageo.2013.06.006. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0098300413001702>.
- Walker, W.E. et al. (2003). 'Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support'. In: *Integrated Assessment* 4.1, pp. 5–17. ISSN: 1389-5176. DOI: 10.1076/iaij.4.1.5.16466. URL: <http://www.tandfonline.com/doi/abs/10.1076/iaij.4.1.5.16466>.
- Webster, Mort, Lisa Jakobovits and James Norton (2008). 'Learning about climate change and implications for near-term policy'. In: *Climatic Change* 89.1-2, pp. 67–85. ISSN: 0165-0009. DOI: 10.1007/s10584-008-9406-0. URL: <http://www.springerlink.com/index/10.1007/s10584-008-9406-0>.
- Webster, Mort, Nidhi Santen and Panos Parpas (2012). 'An approximate dynamic programming framework for modeling global climate policy under decision-dependent uncertainty'. In: *Computa-*

- tional Management Science* 9.3, pp. 339–362. ISSN: 1619-697X. DOI: [10.1007/s10287-012-0147-1](https://doi.org/10.1007/s10287-012-0147-1). URL: <http://www.springerlink.com/index/10.1007/s10287-012-0147-1>.
- Wene, CO (1996). 'Energy-economy analysis: Linking the macroeconomic and systems engineering approaches'. In: *Energy* 21.9, pp. 809–824. URL: <http://www.sciencedirect.com/science/article/pii/0360544296000175>.
- Winkel, Mark et al. (2014). 'Remaking the UK's energy technology innovation system: From the margins to the mainstream'. In: *Energy Policy* 68, pp. 591–602. ISSN: 03014215. DOI: [10.1016/j.enpol.2014.01.009](https://doi.org/10.1016/j.enpol.2014.01.009). URL: <http://dx.doi.org/10.1016/j.enpol.2014.01.009>.
- Wright, Evelyn L. et al. (2010). 'A scenario analysis of investment options for the Cuban power sector using the MARKAL model'. In: *Energy Policy* 38.7, pp. 3342–3355. ISSN: 03014215. DOI: [10.1016/j.enpol.2010.02.005](https://doi.org/10.1016/j.enpol.2010.02.005). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301421510000868>.
- Wynne, B (1992). 'Uncertainty and environmental learning: Reconciling science and policy in the preventive paradigm'. In: *Global Environmental Change* 2.2, pp. 111–127. ISSN: 09593780. DOI: [10.1016/0959-3780\(92\)90017-2](https://doi.org/10.1016/0959-3780(92)90017-2). URL: <http://linkinghub.elsevier.com/retrieve/pii/0959378092900172>.
- Yeh, Sonia et al. (2006). 'An Integrated Assessment of the Impacts of Hydrogen Economy on Transportation, Energy Use, and Air Emissions'. In: *Proceedings of the IEEE* 94.10, pp. 1838–1851. ISSN: 0018-9219. DOI: [10.1109/JPROC.2006.883719](https://doi.org/10.1109/JPROC.2006.883719). URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4016429>.
- Yeh, Sonia et al. (2008). 'Optimizing U.S. Mitigation Strategies for the Light-Duty Transportation Sector: What We Learn from a Bottom-Up Model'. In: *Environmental Science & Technology* 42.22, pp. 8202–8210. ISSN: 0013-936X. DOI: [10.1021/es8005805](https://doi.org/10.1021/es8005805). URL: <http://pubs.acs.org/doi/abs/10.1021/es8005805>.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*".

Final Version as of 20th July 2016 (`classicthesis v2.0.0`).