ROLE OF LOW CARBON ENERGY TECHNOLOGIES IN NEAR TERM ENERGY POLICY

Olaitan P. Olaleye

University of Massachusetts, Amherst, oolaleye@engin.umass.edu

Follow this and additional works at: http://scholarworks.umass.edu/dissertations_2

Part of the Operational Research Commons

Recommended Citation


This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations May 2014 - current by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
ROLE OF LOW CARBON ENERGY TECHNOLOGIES IN NEAR TERM ENERGY POLICY

A Dissertation Presented

By

OLAITAN P. OLALEYE

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2016

Industrial Engineering and Operations Research
ROLE OF LOW CARBON ENERGY TECHNOLOGIES IN NEAR TERM

ENERGY POLICY

A Dissertation Presented

By

OLAITAN P. OLALEYE

Approved as to style and content by:

________________________________
Erin Baker, Chair

________________________________
Hari Balasubramanian, Member

________________________________
Senay Solak, Member

________________________________
Sundar Krishnamurty, Department Head
Mechanical and Industrial Engineering
DEDICATION

To my loving belated mother.

To my new family, Fire and Dayo, for always being there for me.
ACKNOWLEDGEMENTS

I express my gratitude to my advisor, Erin Baker, for her thoughtful guidance and patient mentorship throughout the course of my doctoral education. Her constant dedication was a grounding influence during the varying course of this program and her motivation is the primary reason for this work. I would also like to extend my sincere appreciation to the members of my esteemed committee, Hari Balasubramanian and Senay Solak, for their excellent suggestions and guidance on the dissertation and also for their candid advice regarding a post-graduation career path.

I would also like to recognize the invaluable support and patience of my wife, son and family who always offered motivation and encouragement during times when this task seemed unachievable. Special thanks also to my friends and laboratory mates for their time and support towards making my doctoral studies an awesome experience.

Thanks to the National Science Foundation and the Department of Mechanical and Industrial Engineering at the University of Massachusetts Amherst for providing funding for this different parts of this dissertation. This research was partially supported by NSF award number SES-0745161 and a University of Massachusetts Graduate School Fellowship award. Also, special thanks to the Joint Global Change Research Institute of the Pacific Northwest National Laboratory for providing research support and access to their Evergreen computing cluster.
ABSTRACT

ROLE OF LOW CARBON ENERGY TECHNOLOGIES IN NEAR TERM ENERGY POLICY

FEBRUARY 2016

OLAITAN P. OLALEYE
B.S., [UNIVERSITY OF LAGOS]
M.S., [UNIVERSITY OF MASSACHUSETTS AMHERST]
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by Professor Erin Baker

In the first part of this thesis, we use a multi-model framework to examine a set of possible future energy scenarios resulting from R&D portfolios of Solar, Nuclear, Carbon Capture and Storage (CCS), Bio-Fuels, Bio-Electricity and Batteries for electric transportation. We show that CCS significantly complements Bio-Electricity, while most of the other energy technology pairs are substitutes. From the probabilistic analysis of future energy scenarios we observe that portfolios with CCS tend to stochastically dominate those without CCS; portfolios with only renewables tend to be stochastically dominated by others; and that there are clear decreasing marginal returns to scale. We also find that, with higher damage risk, there is more incentive for technical advancement in CCS and less incentive for development of Solar energy technology.

In the second part of this thesis, we examine the optimal R&D portfolio changes at the different R&D budget levels and how risk in climate damages affects the optimal
R&D portfolio. We find that the optimal portfolio is generally not robust to risk, and the optimal investments in the energy technologies vary with risk in climate damages; however R&D investments in certain energy technologies, such as Nuclear, are robust under the different risk cases. We note that while CCS plays a significant role in the optimal portfolio when there is no risk in climate damages, it plays an even more significant role in the higher climate damage risk cases. We also find that R&D investment in the Biofuels energy technology increases significantly with increase in climate damage risk, while Solar, Batteries for Electric Transportation and Bio-Electricity technologies go out of favor with increases in climate damage risk. We also propose a methodology for obtaining solutions to subset portfolio problems, based on the characteristics of the individual technologies. We prove that the subset portfolio problem is optimal if the individual technology does not interact with any of the other technologies, we confirm this in our empirical portfolio problem.

In the third part of this thesis, we conduct an illustrative global sensitivity analysis on a large scale integrated assessment model with a view to determining the primary drivers of uncertainty in the model and examining the effect of structural uncertainty on the model. We compare our results to a previous paper which conducted a one factor at a time sensitivity analysis and find that both sensitivity methods provide the same result which is different from findings from the previous paper. We find that model interactions are present even in our very limited illustrative analysis. We also conduct most of the steps needed for a full global sensitivity analysis of the model and highlight the challenges in conducting this analysis on the GCAM model. We show that there exist a
need for global sensitivity analysis for accurate determination of the principal drivers of uncertainty in integrated models.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background / Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Key Research Goals</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Thesis Statement and Result Summary</td>
<td>4</td>
</tr>
<tr>
<td>1.4 Thesis Framework Overview</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Thesis Outline</td>
<td>8</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Large Scale Scenario Analysis of Future Low Carbon Energy Options</td>
<td>9</td>
</tr>
<tr>
<td>2.2 R&amp;D Portfolio Optimization</td>
<td>11</td>
</tr>
<tr>
<td>2.3 Uncertainty Analysis in Multi-Model IAMs: Quantification, Modeling and Assessment</td>
<td>12</td>
</tr>
<tr>
<td>3. PART 1: LARGE SCALE SCENARIO ANALYSIS OF FUTURE LOW CARBON ENERGY OPTIONS</td>
<td>15</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>15</td>
</tr>
<tr>
<td>3.1.1 Approach</td>
<td>17</td>
</tr>
<tr>
<td>3.2 Literature Review and Background on Technologies</td>
<td>19</td>
</tr>
<tr>
<td>3.2.1 Literature Review</td>
<td>19</td>
</tr>
<tr>
<td>3.2.2 Background Research on Technologies Modeled</td>
<td>22</td>
</tr>
<tr>
<td>3.3 Methodology</td>
<td>25</td>
</tr>
<tr>
<td>3.3.1 Problem Formulation</td>
<td>26</td>
</tr>
<tr>
<td>3.3.2 Regression Methodology</td>
<td>33</td>
</tr>
<tr>
<td>3.4 Key Results and Analysis</td>
<td>34</td>
</tr>
</tbody>
</table>
3.4.1 Comprehensive Scenario Analysis .................................................. 34
3.4.2 The Timing of Uncertainty Resolution ............................................. 37
3.4.3 Regression Analysis for Substitution and Complementarity Effects Between Technologies .................................................. 38
3.4.4 Probabilistic Scenarios Analysis of Selected Future Energy States ................................................................. 45

3.5 Conclusion .......................................................................................... 52

4. PART 2: R&D ALLOCATION IN A LARGE PORTFOLIO OF 6 CLEAN ENERGY TECHNOLOGIES ........................................................................................................ 56

4.1 Introduction .......................................................................................... 56
  4.1.1 Approach ...................................................................................... 57

4.2 Problem Scope and Solution Approach ................................................. 61
  4.2.1 Problem Scope ............................................................................. 61
  4.2.2 Overview of Modeling Approach .................................................. 62

4.3 Methodology ....................................................................................... 64
  4.3.1 Problem Formulation ................................................................... 64
  4.3.2 Solution Approach: Genetic Algorithm ......................................... 69
  4.3.3 Choosing Budget Levels ............................................................... 74
  4.3.4 Sub-Compartmentalizing the Portfolio Allocation Problem .......... 77

4.4 Results ............................................................................................... 83
  4.4.1 Optimal Portfolio in the No Damage Risk case .............................. 83
  4.4.2 Effect of Climate Damage Risk on the Optimal Portfolio ............. 86
  4.4.3 Sub-compartmentalizing the Portfolio Problem .......................... 92

4.5 Summary/Conclusions ....................................................................... 95

5. PART 3: GLOBAL UNCERTAINTY ANALYSIS OF THE GLOBAL CHANGE ASSESSMENT MODEL: QUANTIFICATION, MODELING AND ASSESSMENT ...................................................................................................... 98

5.1 Introduction ........................................................................................ 98
  5.1.1 Background ................................................................................ 99

5.2 Global Sensitivity Analysis ................................................................ 100
  5.2.1 Global Sensitivity Metrics ............................................................ 100
  5.2.2 Sampling Methodologies ............................................................. 106
  5.2.3 Problem Scope .......................................................................... 111
  5.2.4 Approach ................................................................................... 112

5.3 Methodology: Global Sensitivity Analysis on all Variables of the GCAM Model ........................................................................................................ 113
  5.3.1 GCAM Model Structure ............................................................... 113
  5.3.2 Model Input Parameters ............................................................... 116
  5.3.3 Input Parameter Distribution ...................................................... 116
  5.3.4 Sampling Methodology ............................................................... 116
5.3.5 Model Evaluation ................................................................................. 118

5.4 Methodology: Illustrative Global Sensitivity Analysis study on a Limited Set of Variables in the GCAM Model ................................................................. 119
  5.4.1 Input Parameters Considered ................................................................. 120
  5.4.2 Output Parameter Considered ............................................................... 122
  5.4.3 Input Parameter Distribution and Sampling Methodology ............... 122
  5.4.4 Global Sensitivity Analysis Methodology ........................................... 126
  5.4.5 One Factor at a Time (OFAT) Sensitivity Analysis ......................... 126

5.5 Results ........................................................................................................ 127
  5.5.1 Global Sensitivity Analysis of the GCAM model ......................... 127
  5.5.2 Illustration Global Sensitivity Analysis on a Selection of Variables ..................................................................................................................... 129

5.6 Conclusion .................................................................................................... 132

APPENDICES

A. R&D FUNDED TECHNOLOGIES, SUB-TECHNOLOGIES, FUNDING LEVELS, SUCCESS PROBABILITIES AND R&D ENDPOINTS .................. 138

B. INTEGRATING TECHNICAL CHANGE ......................................................... 140

C. ENERGY SCENARIOS AND R&D OUTCOMES ......................................... 141

D. UNCERTAINTY IN CLIMATE DAMAGES .................................................. 142

E. EVALUATING ENERGY SCENARIO PROBABILITY ................................... 143

F. OPTIMAL R&D PORTFOLIO ................................................................. 145

G. CONSISTENCY OF THE SOLUTION .......................................................... 148

H. SUB-COMPARTMENTALIZATION .............................................................. 149

I. BIO-FUELS TECHNOLOGICAL PATHS ..................................................... 153

J. GLOBAL SENSITIVITY ANALYSIS METHODS ......................................... 164

BIBLIOGRAPHY .......................................................................................... 168
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Parameters and variables of the scenario evaluation model.</td>
<td>28</td>
</tr>
<tr>
<td>2: Damage Uncertainty Cases.</td>
<td>33</td>
</tr>
<tr>
<td>3: Regression results for all uncertainty cases and statistically significant technology pairs.</td>
<td>39</td>
</tr>
<tr>
<td>4: An illustration of the R&amp;D effectiveness of the sub-technologies.</td>
<td>43</td>
</tr>
<tr>
<td>5: Characteristics of selected portfolios.</td>
<td>49</td>
</tr>
<tr>
<td>6: Parameters of the Genetic Algorithm</td>
<td>72</td>
</tr>
<tr>
<td>7: R&amp;D effectiveness of the sub-technologies</td>
<td>76</td>
</tr>
<tr>
<td>A.1: R&amp;D funding, success probabilities and cost endpoints for all technologies and sub-technologies considered.</td>
<td>138</td>
</tr>
<tr>
<td>C.1: Scenarios and Outcomes Definition and Calculation</td>
<td>141</td>
</tr>
<tr>
<td>D.1: Damage Risk Cases. The table shows the percentage GDP loss for a 2°C rise in mean atmospheric temperature for each of the risk cases.</td>
<td>142</td>
</tr>
<tr>
<td>F.1: Optimal Portfolio for the No Damage Risk Case ($ million R&amp;D).</td>
<td>146</td>
</tr>
<tr>
<td>I.1: Bio-Fuels Funding Trajectories.</td>
<td>154</td>
</tr>
<tr>
<td>I.2: Aggregated Probabilities for all Bio-Fuels Paths.</td>
<td>155</td>
</tr>
<tr>
<td>I.3: Bio-Fuels funding levels, success probabilities and cost endpoints.</td>
<td>163</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Flow diagram for scenario analysis study.</td>
<td>18</td>
</tr>
<tr>
<td>2: Marginal abatement cost measured as percentage of gross domestic product GDP.</td>
<td>32</td>
</tr>
<tr>
<td>3: Expected societal utility of energy scenarios ordered by increasing utility in the No Uncertainty case.</td>
<td>36</td>
</tr>
<tr>
<td>4: Effect of early or late resolution of climate damages.</td>
<td>38</td>
</tr>
<tr>
<td>5: Regression results for statistically significant technologies and technology pairs.</td>
<td>40</td>
</tr>
<tr>
<td>6: Close-up of the abatement cost curves of Solar, Batteries for Electric Transportation and Bio-fuels technologies for the year 2050.</td>
<td>41</td>
</tr>
<tr>
<td>7: Illustration of the R&amp;D effectiveness of the sub-technologies.</td>
<td>44</td>
</tr>
<tr>
<td>8a and 8b: CDF's of some selected energy scenarios in the No Uncertainty (8a) and High Uncertainty (8b) cases.</td>
<td>51</td>
</tr>
<tr>
<td>9: R&amp;D Portfolio Solution Overview.</td>
<td>64</td>
</tr>
<tr>
<td>10: Optimal R&amp;D Portfolio to Utility chart for the $300 million to $6.1 billion R&amp;D budget levels when there is no uncertainty in climate damages.</td>
<td>85</td>
</tr>
<tr>
<td>11: Optimal R&amp;D Portfolio to Utility chart across all the R&amp;D budget levels when there is no uncertainty in climate damages.</td>
<td>85</td>
</tr>
<tr>
<td>12: Societal Utility by Damage Risk Cases for all the R&amp;D budget levels.</td>
<td>88</td>
</tr>
<tr>
<td>13: Change in Optimal Portfolio with Damage Risk with respect to the No Damage Risk</td>
<td>90</td>
</tr>
<tr>
<td>14: Abatement Cost curves with only advancement in one sub-technology.</td>
<td>91</td>
</tr>
<tr>
<td>15: Marginal abatement cost curves 60% - 100%.</td>
<td>91</td>
</tr>
<tr>
<td>16: Change in optimal portfolio with respect to CCS.</td>
<td>95</td>
</tr>
<tr>
<td>17: Pseudo-Random, Quasi-Random Sobol sequence and Scrambled Quasi-Random Sobol sequence generated samples in respective order.</td>
<td>108</td>
</tr>
</tbody>
</table>
18: Hierarchical tree structure of the base configuration file in the GCAM model. 
Child nodes are in blue and roots in green. ......................................................... 115

19: Resulting Model Input and Output Distribution. ............................................. 125

20: Three Dimensional Plot of the Input Parameter Space Distribution. .............. 126

21: Normalized OFAT and Global Sensitivity Analysis GSA analysis metrics. ....... 131

F.1: Optimal R&D Portfolio to Utility chart across all the R&D budget levels when 
there is Medium uncertainty in climate damages. ............................................... 147

F.2: Optimal R&D Portfolio to Utility chart across all the R&D budget levels when 
there is High uncertainty in climate damages. .................................................... 147

G.1: Variability of the genetic algorithm solution. ............................................... 148

H.1: Change in optimal portfolio with respect to Nuclear. ..................................... 150

H.2: Change in optimal portfolio with respect to Bio-Electricity. ......................... 151

H.3: Change in optimal portfolio for some selected budget levels when CCS and 
Bio-Electricity is excluded. .................................................................................. 152

I.1: Bio-Fuels technological paths from feedstock to end product. ....................... 154

I.2: Pre-defined Bio-Fuels Technological Paths. ................................................... 157

I.3: STP1 Best Budgets to Weighted Expected Utility. ........................................... 159

I.4: STP2 Best Budgets to Weighted Expected Utility. ........................................... 159

I.5: Hydrolysis sub-technology Best Budgets to Weighted Expected Utility. .......... 161

I.6: Gasification sub-technology Best Budgets to Weighted Expected Utility. ........ 162

J.1: Cumulative distribution function of the labor productivity growth relative to 
radiative forcing using the distance to CDF based Kolmogorov Smirnov test. ...... 164

J.2: Cumulative distribution function of the income elasticity of demand relative to 
radiative forcing using the distance to CDF based Kolmogorov Smirnov test. ...... 165

J.3: Cumulative distribution function of the AEII relative to radiative forcing using 
the distance to CDF based Kolmogorov Smirnov test. ........................................ 165

J.4: Density function of the conditional labor productivity growth relative to 
radiative forcing using the density based Borgonovo Importance measure. ........ 166
J.5: Density function of the conditional income elasticity of demand relative to radiative forcing using the density based Borgonovo Importance measure. ............ 166

J.6: Density function of the conditional AEII relative to radiative forcing using the density based Borgonovo Importance measure. ......................................................... 167
CHAPTER 1

INTRODUCTION

The general context of this PhD study revolves around the role of low carbon energy technologies in near term energy policy planning to address climate change. We explicitly explore the impact of having technological advancement in a portfolio of six (Solar, Nuclear, CCS, Bio-Fuels, Bio-Electricity and Batteries for electric transportation) clean energy technologies. Our goal is to ease the process of decision making by providing clear actionable insights on current decisions that can mitigate potential damages from climate change.

1.1 Background / Motivation

Humanity, over the course of its evolution, has overcome several severe challenges. Few adversities, however, have been both potentially calamitous, and causal, as climate change. It is simply the most exigent issue humanity faces today. Its severity makes it necessary for different approaches to be taken to mitigate its impact. One such approach is accelerating the rate of technological development in energy technologies through research and development (R&D).

Directed R&D interventions have led to some of the major improvements in man’s quality of living, e.g. the steam turbine. Such efforts highlight the role of deliberated R&D quests in developing promising energy technologies to address the climate change debacle. As such, while the Department of Energy (DOE) expends billions annually in R&D funding of energy technologies (Anadon et al., 2015), these funds are limited and competition exists amongst this ever changing group of energy
technologies. The R&D process is also inherently uncertain; expert elicitation studies are one way to assess the viability of these technologies. Another major source of uncertainty is the severity of the resulting climate damages. There exists a considerable paucity of information on both the future level of climate damages and the risk of occurrence of such damages. These factors make addressing the optimal allocation of R&D funds to futuristic clean energy technologies a crucial and rather interesting area of research.

As such, technology policy should account for the different future realizations of climate damages as well as for the uncertainty in the R&D process. It is in this light that we examine the optimal R&D technology policy to mitigate the impact of climate change. We consider major low-carbon energy technologies that are capable of being deployed on a large scale and that have the potential to lead to substantial reduction in energy cost. We therefore consider technological advancements in Solar Photovoltaic, Nuclear, Carbon Capture and Storage (CCS), Liquid Bio-Fuels, Electricity from Biomass, and Batteries for transportation. This set of technologies was chosen for three reasons. First, there was a set of consistent expert elicitation studies available on this set of technologies. Second, these technologies have the potential for significant clean energy generation. Third, R&D in these technologies has the potential to result in significant cost reduction, relative to learning by doing on its own.

Here we point out the specific strengths of each of these technologies. The earth’s solar resource has the potential to produce several orders of magnitude more energy than our current total world energy consumption (Perez and Perez, 2009). CCS has the ability to curtail climate emissions in a short period without significantly changing the current energy infrastructure (Metz et al., 2005). Nuclear energy, while controversial from a
security and safety perspective, has been shown to have the lowest carbon footprint of all the current baseload energy technologies (Edenhofer et al., 2014). Bio-Electricity and Liquid Biofuels were selected given the vast amount of the biomass resource in the environment, and hence their potential for renewable energy generation. Batteries for Electric Transportation are a potential solution to the tremendous challenge that the transportation sector poses to carbon mitigation. We note that while this set of energy technologies does not constitute all promising clean energy technologies, they were readily analyzable at the time of the elicitation studies. We concede that the flux of current promising technologies may be ever-changing, hence the value of this study might not be in the specific set of results but in the methodology used. The energy technologies and the sub-technologies considered within each technology are discussed in further detail in Section 3.2.2 and Appendix A.

1.2 **Key Research Goals**

The key goals of this thesis are highlighted below.

- To characterize the relationship between different energy technologies based on their benefit to the society.
  - To explicitly assess technology pairs and groups to determine whether the technologies are substitutes or complements.
  - To evaluate how the substitutive or complementary behaviors change with uncertainty in climate damages.

- To examine the risk profiles of different R&D technology portfolios.
  - To examine for stochastic dominance relationships between these portfolios.
To evaluate how the stochastic dominance relationships change with uncertainty in climate damages.

- To determine the optimal allocation of R&D funds to the technology portfolio based on some initially assessed probability distributions over their energy costs.
- To determine, given a specific energy R&D investment budget, the optimal energy portfolio.
- To determine if and how the optimal energy portfolio changes with uncertainty in climate damages.

- To investigate global sensitivity analysis and its effects in a large scale IAM framework.
  - To provide a framework that allows us to model uncertainty propagation through an integrated assessment model (IAM).
    - To approach this from a global uncertainty analysis based simulation perspective.

1.3 Thesis Statement and Result Summary

This thesis is divided into three main chapters. The various chapters address different aspects of the R&D technology policy response to climate change.

In the first part, Chapter 3, we conduct a comprehensive scenario analysis of low carbon energy technologies. We do this by exploring all possible energy states that can result in the future, if R&D in these prospective technologies is conducted.

One of the results from this first part of the thesis is that there is a need for directed R&D analysis. This is because certain R&D budget allocations (portfolios) are clearly dominated by others, even much smaller allocations. Specifically we observe that
R&D portfolios without CCS are stochastically dominated by portfolios with R&D development in CCS. We also find that CCS complements Bio-Electricity and that most of the other energy technology pairs are significant technological substitutes. Finally, we find that, as the risk of climate damages increases, there is more incentive for R&D investment in the CCS and Bio-Fuels energy technologies.

In the second part of the thesis, Chapter 4, we examine the optimal allocation of R&D funding amongst the competing technologies. We do this in the face of uncertainty in the level of climate damages and the R&D technological process.

One of the results from this section is that the relationships observed between energy technologies in the first chapter provide insights into the optimal portfolio structure. We find that the optimal portfolio makes R&D investment in CCS, In-organic solar cells, and the Light Water Reactor Nuclear sub-technologies a priority. We discover that the optimal portfolio varies with uncertainty in climate damages as we see that optimal R&D investment in CCS and Bio-Fuels increases as the risk of climate damages increases. We find that the variation in the optimal portfolio across the different future climate damage scenarios is due to the carbon abatement characteristics of the energy technologies under these scenarios, and the complementary or substitutive relationships between the energy technologies. We also show that the R&D portfolio allocation problem can be decomposed into smaller sub-problems when no interactions exist between the energy technologies. Based on a regression from Chapter 3 to assess technology interactions, we apply the decomposition methodology to our R&D portfolio problem. We find that, when technologies that have no significant interaction with other
technologies are excluded at their optimal funding level, the resulting sub-problem has the same optimal portfolio with the large R&D portfolio.

In the third part of this thesis, Chapter 5, we provide a framework for conducting global uncertainty analysis of a large scale integrated assessment model (IAM), the Global Change Assessment Model (GCAM). We also conduct an illustrative global uncertainty analysis of the GCAM model, based on the main parameters from a prior local uncertainty analysis study.

One of our results from this section is that given the large number of input parameters in the model, full scale global uncertainty analysis of the model is technically impossible, as the number of the model inputs is more than current computational resources allow. We also find that when all model inputs are varied simultaneously, the model does not solve.

From the illustrative global sensitivity analysis, our results include: that the importance ranking of the model inputs, from the global sensitivity study, is different from that of the previous local sensitivity study. This is because both studies use slightly different models. We however find that when global and local sensitivity analysis is conducted on the selected parameters in GCAM, we obtain the same ranking of parameters.

1.4 Thesis Framework Overview

In this section, we provide a brief overview of the general problem framework and highlight the methodology on which each of the different studies conducted is based. We provide specific details of our research methodology in the approach section of each of the subsequent chapters.
To conduct the various studies discussed in the thesis statement section above, we utilize a multi-model framework comprising the Global Change Assessment Model (GCAM), a stochastically reformulated version of the Dynamic Integrated Model of Climate and the Economy (DICE), and a portfolio optimization model. The GCAM model is a technologically-detailed integrated assessment model, allowing us to model the different mixes of futuristic energy technologies. The DICE model is a relatively computationally inexpensive integrated assessment model, which enables us to conduct large scale scenario analysis, while incorporating the dynamic impacts on social utility. Our portfolio optimization model is an R&D allocation model, which allows the integration of technological uncertainty into the framework.

This framework enables us to address several objectives. The GCAM model provides a structure to model all possible future energy states that can result from advancement in the different energy technologies. The DICE model integrates these energy states and enables us to examine whether the different energy technologies are substitutes or complements. We can also examine how these relationships between the technologies, substitutes or complements, affect the overall contribution of the technologies to societal utility. It also enables us to evaluate the societal utility, of any of the possible future energy states that can result from having R&D investment in the technologies.

In our portfolio optimization model, we integrate probabilistic distributions over the energy scenarios based on previously performed expert elicitation, for specific R&D funding levels, at different R&D target specifications. This enables us to conduct stochastic dominance relations between the different possible R&D allocations. Finally,
our portfolio optimization model enables us to obtain the optimal allocation of R&D funding to the technologies. Additionally, the portfolio model allows us to explore how the optimal R&D allocations vary with uncertainty in climate damages.

For the third part of the thesis, we conduct a global uncertainty analysis on the GCAM model with a view to investigating how uncertainty in the model output is explained by uncertainty in the model inputs. Using a quasi-random low discrepancy method, the Sobol sequence, we generate draws from the input distribution space. We evaluate the GCAM model for each set of the generated samples and evaluate the model outputs based on variance decomposition, density and distribution based global sensitivity methods.

1.5 Thesis Outline

The subsequent portions of this thesis are organized as follows: Chapter 2 gives a review of the background research that led to this work, as well as a review of the relevant literature. Chapter 3 provides details on the first part of this thesis. It focuses on the global scenario analysis of clean energy technologies under the different possible climate states. Chapter 4 is focused on the portfolio optimization approach to the R&D technology policy response to climate change problem. Chapter 5 is dedicated to the global uncertainty analysis of the GCAM model. Chapter 6 concludes the thesis, with possible recommendations for future work. In the appendix, we provide supplementary content including the re-modification of the Bio-Fuels technological paths. The re-modification was conducted so that the technological paths can be independent of each other and thus fit within the structure of our portfolio optimization model.
CHAPTER 2
LITERATURE REVIEW

2.1 Large Scale Scenario Analysis of Future Low Carbon Energy Options

One approach to addressing the impact of R&D on climate change is scenario analysis (Kobos et al., 2006; Clarke et al., 2008; Edelhofer et al., 2010; McJeon et al., 2011; Pugh et al., 2011; Luderer et al., 2012; Shell Group, 2005). Scenario analysis entails the characterization and evaluation of internally coherent future energy states of the world that result from certain underlying presumptions about the initial states (Huss, 1988; Kahn and Wiener, 1967; Swarta et al., 2004).

With scenario analysis, one has the choice to selectively (e.g. (Yohe 1991; Nakicenovic, . et al. 2000)) or comprehensively (McJeon, Clarke et al. 2011) assess the resulting possible states of the world. Morgan & Keith (2008) show that selective scenario analysis leads to ‘systematic overconfidence’ as this causes the decision analyst to focus only on the scenarios modeled, and ignore possible extreme events that are not represented. Additionally the fact that most previous energy forecasts have been inaccurate (e.g. (Lovins, 1976; DOE, 1979; Craig et al., 2002; Smil, 2003; Kirsch, 2005)) emphasizes the need to consider all possible outcomes.

Additionally, when using scenario analysis for decision analysis, a choice exists on whether to analyze scenarios deterministically (e.g. (Nakicenovic, . et al. 2000), (McJeon, Clarke et al. 2011)) or probabilistically (e.g. (O'Neill 2004), (Pugh, Clarke et al. 2011)). Probabilistic scenario analysis entails assigning a probability distribution over the scenarios. Probabilistic comprehensive scenario analysis therefore has the advantage that
it allows the evaluation of the full space of the scenarios within a consistent framework (e.g. (Groves and Lempert 2007), (Schneider 2001)). On the other hand, the use of probabilistic scenario analysis has been faulted as inherently subjective ((Grübler and Nakicenovic 2001), (Schneider 2001)) through the assessment of the probability distributions and possibly overly cumbersome.

Other approaches to R&D decision analysis exist including, but not limited to, sensitivity analysis (e.g. (Dowlatabadi 1998)), optimal portfolio analysis ((Blanford and Weyant 2005), (Bosetti, Carraro et al. 2009), (Blanford 2009), (Baker and Solak 2011), (Diaz, Bunn et al. 2011), (Baker and Solak 2013)) and extreme space estimation (Moss, Edmonds et al. 2010). Recent work building on expert judgments includes Anadon et al. (2015), in which diverse expert elicitations are being harmonized and aggregated. Such pooling of diverse opinions is a great tool for characterizing uncertainty, however, the role that technologies play in the economy is a key to fully characterizing possible R&D outcomes.

*Energy Technology Interactions: Substitutes or Complements.* A few studies (e.g. (Chow, Kopp et al. 2003, Edenhofer, Knopf et al. 2010, McJeon, Clarke et al. 2011)) have noted the existence of dependencies between the gains or cost reductions from having advancement in energy technologies; however, no paper that we know of has developed a framework to quantitatively assess the degree and nature of these relations within an IAM framework.
2.2 R&D Portfolio Optimization

Another approach to addressing the impact of technological R&D investment in climate change mitigation is R&D portfolio optimization analysis. Numerous studies have investigated the optimal allocation of R&D funds to energy technologies in the face of climate change (Eilat et al., 2006; Baker and Solak, 2011; Blanford and Weyant, 2005; Blanford, 2009; Bosetti et al., 2009; Baker et al., 2014). A subset of the numerous approaches to optimal R&D portfolio allocation include the following: probability-based portfolio allocation approaches [non-parametric expert opinion based approaches (Baker and Solak, 2014; Baker and Solak, 2011; Bosetti et al., 2009); parametric-based approaches (Blanford, 2009)], real option valuation approaches [fuzzy logic based approaches ; (Wang and Hwang, 2007; Bardhan et al., 2006)], multi-criteria portfolio evaluation approaches (Eilat et al., 2006; Linton et al., 2002).


Portfolio Decomposition: Another focus of the R&D portfolio allocation chapter is the question of efficient decomposition of R&D portfolio allocation problems. While these constrained knapsack optimization problems are known to be NP-complete (e.g. (Garey & Johnson, 1979)), numerous methods exist for reducing the complexity of these knapsack problems, including dominance relation techniques (Andonov et al., 2000; Poirriez et al., 2009) when strict dominance conditions exists between the constituent
items. We show in (Olaleye and Baker, 2015) that while such dominance relations can exist in some cases between some energy portfolios in our 6 energy technology problem, such orderings are not universal and may not always hold across all the R&D portfolios. To exhaustively evaluate all possible portfolios is computationally at least as intensive as exhaustively solving the portfolio problem (Olaleye and Baker, 2015). We note that our particular problem has the characteristic that the sub-technologies are not independent of each other, hence stochastic dominance techniques cannot be used across all portfolios. We therefore approach this problem from a genetic algorithm optimization perspective given the scale and the nature of the problem (e.g. Hassan, Cohanim et al. 2005).

2.3 Uncertainty Analysis in Multi-Model IAMs: Quantification, Modeling and Assessment

Integrated assessment models (IAMs) are a valuable source for aiding decision making regarding climate change (e.g. (Parson, 1994; Morgan and Henrion, 1990)). As observed from the various IPCC reports (e.g. IPCC4 2007), IAMs have been particularly useful in providing a consistent framework where numerous simulations of the future, with different initial model input assumptions, can be carried out. As these IAMs are, by nature, large interacting models with projections running centuries later (e.g. (Nordhaus 2008),(Calvin, Clarke et al. 2011)), it is of utmost importance that the base inputs to these models are as accurate as possible.

The structure of these IAM models also usually requires several inputs which are subjective (e.g. (van der Sluijs, 2002)). These inputs interact with one another leading usually to accentuation (e.g. (Draper, 1995)) of the uncertainty therein or rarely dampening (van der Sluijs, 1996) of such uncertainties. While most of the previous
studies on uncertainty analysis have focused on one input factor at a time uncertainty analysis (e.g. (Scott et al., 1999)), several authors have noted the deficiency of the technique as it does not capture the uncertainty due to model structure and parameter dependence. Satelli and Campalongo (2000) and Jing Yang (Yang, 2011) have noted that most studies examining the propagation of uncertainties through IAMs should primarily avoid ‘one input factor at a time’ analysis, given such analysis will be biased as uncertainty in the other input parameters is not considered simultaneously. Most new studies have therefore focused on global sensitivity approaches to examining uncertainties about the model inputs (e.g. (Anderson et al., 2014)). These however are not the only source of uncertainty in IAM models. As Draper et al. (1987) and van der Sluijs (1996) noted, aside such parametric uncertainty about the inputs, other sources of uncertainty exist, some of which include uncertainties relating to the structure of the IAM model and the stochasticity of the model outputs (e.g. (Draper et al., 1987; van der Sluijs, 1996; Kann and Weyant, 2000; Golub et al., 2014)). This study is primarily aimed at studying parametric uncertainty propagation introduced through model inputs from a global sensitivity perspective.

Numerous studies have also used different methods to assess the propagation of these parametric uncertainties caused by model inputs in IAMs. Helton & Davis (2003) provide a good review of these techniques; they include Monte Carlo sampling (e.g. (Hoffman and Hammonds, 1994; Webster et al., 2002)), probability theory (e.g. Ferson and Ginzburg 1996), differential analysis (e.g. (Tomovic and Vukobratovic, 1972; Lewins and Becker, 1982)), response surface methodology (e.g. (Kleijnen et al., 1992)), Sobol variance decomposition (e.g.(Sobol’, 1993)) and Fourier amplitude sensitivity test
Some studies have also investigated uncertainty propagation from a theoretical perspective (e.g. (Draper, 1993; Chatfield, 1995; Draper, 1995)) through various means such as Bayesian updating.

A few other studies have also examined the accuracy of multi-model ensembles (Reichler and Kim, 2008; Barnston et al., 2003; Doblas-Reyes et al., 2005). Barnston et al. (2003) show that their multi-model ensemble leads to bias reduction, as the models exhibit differing characteristics. Doblas-Reyes et al. (Doblas-Reyes et al., 2005) find similarly that a multi-model approach is valuable.

We note that, while the literature is fairly extensive on global sensitivity analysis in decision models, little or no literature exists on extending this methodology to large scale models due to the data and computationally intensive nature of such studies.

Finally, this thesis relies primarily on two papers in the uncertainty analysis sphere: Anderson et al. (2014), who conduct a global sensitivity of the DICE model using the variance decomposition method, and Scott et al. (1999), who conducted a one factor at a time (OFAT) uncertainty analysis of the MiniCam model, a predecessor to the GCAM model.
CHAPTER 3
PART 1: LARGE SCALE SCENARIO ANALYSIS OF FUTURE LOW CARBON ENERGY OPTIONS¹

3.1 Introduction

Scientists largely agree that man’s actions – past and present – are causing the earth to warm up (IPCC AR4 et al., 2007). Uncertainty still exists, however, about the severity of the resulting climate damages and on the future of technological change in energy technologies. Technology policy, therefore, should account for different future realizations of climate damages and technical change. It is in this light that we evaluate potential technological advancements in six major low-carbon energy technologies, including Solar Photovoltaic, Nuclear, Carbon Capture and Storage, Liquid Bio-Fuels, Electricity from Biomass, and Batteries for Electric Transportation.

In contrast to a previous study that investigated the role of technical change on climate policy through large scale scenario analysis (McJeon, Clarke et al. 2011), the crux of our research is to explore the role of climate damage uncertainty on the relative impacts of the different energy scenarios, where a scenario is one possible energy future, represented by a set of cost and performance parameters over the six technologies. Our approach, which uses multiple models of differing levels of complexity and builds on the results of expert elicitations, allows us to study previously unaddressed questions on the

¹ This chapter of the thesis is published in the Energy Economics Journal (Olaleye and Baker, 2015) and is reprinted with copyright clearance permission from the journal. Minor repetitions in the description of the technologies and the literature review exist as they are elaborated on in later sections.
relationships between energy technologies, and to rank future energy scenarios in terms of social utility.

In order to evaluate the likelihood of different possible future energy scenarios, we turn to previously performed expert elicitations. These studies follow an explicit protocol in order to elicit subjective probabilities over energy futures from a wide range of scientists and engineers (for a summary see Baker et al. (2014)). Hence, the results from these studies are inherently subjective. However, as a number of panels and studies have pointed out, expert elicitations are often the best way to characterize future uncertainty over events such as future technological breakthroughs [e.g. (Boring, et al., 2005), (Mastrandrea, et al., 2010)].

Combining expert judgments with multiple models allows us to evaluate the expected social welfare of different energy technology research and development (R&D) portfolios, where a portfolio is a set of particular funding levels for each technology, and is associated with a probability distribution over scenarios. By combining probability distributions derived from the expert elicitations with the economic outcomes of technological advancement derived from economic models, we are able to evaluate stochastic dominance relations between different energy portfolios.

The central theme of this paper is to conduct a scenario analysis of promising future energy technologies with a view to aiding near term energy policy decision making. To do this, we address the following specific study goals: One is to understand the relative importance of advancement in individual technologies in an economy facing uncertain climate damages. Another goal is to understand the interactions between pairs of advanced energy technologies in the economy, and how these interactions change with
uncertainty in climate damages. A third goal is to examine stochastic dominance between R&D funding portfolios, and to understand how uncertainty in climate damages affects these stochastic dominance relations.

3.1.1 Approach

To address the questions raised above, two integrated assessment models (IAMs) are used, the Global Change Assessment Model (GCAM) and a stochastic version of the Dynamic Integrated Model of Climate and the Economy (DICE). The GCAM model is technologically detailed, allowing us to model the different mixes of futuristic energy technologies. The stochastically reformulated DICE model is computationally inexpensive, enabling large scale scenario analysis, while incorporating the dynamic impacts on social utility and decision making under uncertainty about climate damages. This approach enables us to examine how dependencies between technologies affect the overall benefits of having such energy technologies in our R&D portfolio when climate damages are uncertain, and to determine dominance relationships between different energy portfolios.

Our specific approach is as follows. We generate a large set of energy technology scenarios, encompassing combinations of price and performance parameters for our six technologies. These scenarios are first run through the technologically-detailed GCAM model, under a series of different carbon taxes, in order to estimate the impact of technological change on the cost of reducing carbon emissions.\(^2\) We then use these

\(^2\)The term “carbon emissions” here actually refers to the CO\(_2\)– equivalent of the set of all other greenhouse gases, as given in Van Vuuren (2008)
estimated MACs to implement technological change into the DICE model. Our stochastic version of the DICE model includes uncertainty and learning about climate damages, and can calculate an expected utility associated with each energy scenario. This approach is shown in Figure 1.

![Flow diagram for scenario analysis study.](image)

**Figure 1: Flow diagram for scenario analysis study.**

*Rectangles represent models, elongated semi circles represent processes and rounded rectangles represent outputs.*

To understand the interplay between the different energy technologies, we conduct a simple regression analysis over the set of technology scenario outputs from the DICE model. The independent variables represent the level of technological advancement. The effect of the technologies on the resulting dependent variable, the expected utility, is then evaluated through the regression.

Finally, drawing on the previously performed expert elicitations, we assign probability distributions over the set of technology scenarios, conditional on the specific R&D portfolio to obtain dominance relations.

The remainder of this paper is organized as follows: Section 3.2 provides a review of the literature and the background research leading to this work. In Section 3.3, we present the problem formulation, the models used in the study, our calibration of these
models, the regression technique, and the calculation of probabilities. In Section 3.4 we present the results while Section 3.5 concludes the paper and gives future research recommendations.

3.2 Literature Review and Background on Technologies

3.2.1 Literature Review


With scenario analysis, one has the choice to selectively, based on the purpose of the study and feasibility (e.g. Yohe 1991, Nakicenovic, . et al. 2000), or comprehensively (McJeon, Clarke et al. 2011) assess the resulting possible states of the world. Morgan & Keith (2008) show that selective scenario analysis leads to ‘systematic overconfidence’ as this causes the decision analyst to focus only on the scenarios modeled, and ignore possible extreme events that are not represented. Additionally the fact that most previous energy forecasts have been inaccurate (e.g. Lovins 1976, DOE 1979, Craig, Gadgil et al. 2002, Smil 2003, Kirsch 2005) emphasizes the importance of considering all possible outcomes. We use comprehensive scenario analysis, conditioned on the data available from the expert elicitations.
When using scenario analysis for decision analysis, a second choice exists, on whether to analyze scenarios deterministically (e.g. Nakicenovic, et al. 2000, McJeon, Clarke et al. 2011) or probabilistically (e.g. O'Neill 2004, Pugh, Clarke et al. 2011). Probabilistic scenario analysis entails assigning a probability distribution over the scenarios. Probabilistic comprehensive scenario analysis therefore has the advantage that it allows the relative evaluation of the full space of the scenarios in finite cases and the determination of the distribution over these scenarios within a consistent framework (e.g. Groves and Lempert 2007, Schneider 2001). On the other hand, the use of probabilistic scenario analysis has been faulted as inherently subjective (Grübler and Nakicenovic 2001, Schneider 2001) through the assessment of the probability distributions and possibly overly cumbersome. In this paper we present both a global deterministic scenario analysis and a probabilistic portfolio analysis.

Other approaches to R&D decision analysis exist including sensitivity analysis (e.g. (Dowlatabadi 1998)), optimal portfolio analysis (Blanford and Weyant 2005, Bosetti, Carraro et al. 2009, Blanford 2009, Baker and Solak 2011, Diaz, Bunn et al. 2011, Baker and Solak 2013) and extreme space estimation (Moss, Edmonds et al. 2010). Recent work on expert judgments by Anadon, Baker et al. (Under Review), in which diverse expert elicitations are being harmonized and aggregated, also complement the probabilistic scenario analysis, optimal portfolio analysis and the extreme space estimation approaches. The study (Anadon, Baker et al. Under Review) noted that pooling of diverse opinions is a useful tool for characterizing uncertainty. They also note that technology interactions with each other and the economy play a significant role in characterizing the impact of R&D.
Though a few studies (e.g. (Chow, Kopp et al. 2003, Edenhofer, Knopf et al. 2010, McJeon, Clarke et al. 2011)) have noted the existence of dependencies between the gains or cost reductions from advancement in energy technologies, no paper that we know of has developed a framework to quantitatively assess the degree and nature of these relations within an IAM framework. This study falls in the category of comprehensive probabilistic scenario analysis, where the uncertainties are in climate damages and technological outcomes. We rely on previous expert elicitations by Baker et al. (Baker, Chon et al. 2008, Baker, Chon et al. 2008, Baker, Chon et al. 2009, Baker, Chon et al. 2009, Baker, Chon et al. 2010, Baker and Keisler 2011) to serve as inputs to assess the likelihood of technological development given R&D investment in the different technologies.

As several previous studies have noted, research using IAMs has some inherent limitations: the input technology characteristics have to be estimated as they are usually not well known [e.g. (Baker & Solak, 2013)]; there exist significant questions on the appropriate methodology and time frame, for resolution of climate damages and mitigation of uncertainty [e.g. (Grübler & Messner, 1998), (Weyant & Olavson, 1999), (Webster M., 2002), (Golub, Narita, & Schmidt, Uncertainty in integrated assessment models of climate change: alternative analytical approaches, 2014)] and model bias and knowledge incompleteness [e.g. (Risbey, Kandlikar, & Patwardhan, 1996)]. Nevertheless, IAMs have proven to be a useful tool for gaining insights and informing policy (Kunreuther, et al., 2014).

Another issue that has been discussed in the literature is uncertainty resolution. Most studies have focused on a two stage model for uncertainty resolution as this eases
computational complexity considerably [e.g. (Yohe, Andronova, & Schlesinger, 2004), (Webster M., 2008.)]. A few studies have used multi-stage models, but have had to simplify the IAM due to the computational cost [e.g. (Webster, Santen, & Parpas, 2012), (Crost & Traeger, 2012), (Kelly & Kolstad, 1999)]. Some results indicate, however, that many of the insights can be gained by using a 2-stage model with perfect learning to approximate multi-stage models with partial learning [ (Baker E., 2005), (Webster, Santen, & Parpas, 2012)]. We use a two-stage model in this paper.

3.2.2 Background Research on Technologies Modeled

We consider an R&D portfolio consisting of six low carbon energy technologies, including Carbon Capture and Storage (CCS), Nuclear fission, Solar Photovoltaics, Bio-Electricity, Liquid Biofuels, and Batteries for Electric Transportation. These include many of the key technological advancements capable of offering notable improvements from current energy costs (Baker and Solak 2011). We note that this selection of energy technologies is not comprehensive. It is based on a set of existing expert elicitation studies that were designed explicitly to be combined with GCAM. Most notable is the absence of wind energy. While this is a promising technology, to date there have been no expert elicitation studies on this topic. Ideally future elicitation studies will include wind and other technologies. A brief discussion of each technology and their sub-technologies is given below and summarized in the appendix in Table A.1, including a graphical
summary of the six energy technologies, their 17 sub-technologies and R&D cost endpoints.\(^3\)

**Carbon Capture and Storage:** This technology captures carbon emissions from emitting plants (primarily electricity generation in this paper) and allows them to be stored in underground aquifers. Three different sub-technologies are considered: *Pre-Combustion* in association with Integrated Gasified Combined Cycle generation, *Chemical Looping*, and different types of *Post Combustion* (Baker, Chon et al. 2009).

**Nuclear:** We considered technological advancement in large-scale Nuclear Fission technology, including *Advanced Light Water Reactors, High Temperature Reactors* and *Fast Reactors* (Baker, Chon et al. 2008).

**Solar Photovoltaic:** R&D development is considered in *Purely Organic Solar Cells, In-organic Solar Cells* and *Third generation* (including multi-junction concepts and quantum dots) solar cell technologies (Baker, Chon et al. 2009).

**Bio-Electricity:** This technology encompasses different techniques for the generation of electricity from cellulosic biomass feedstock. While the technology has a relatively inexpensive feedstock, it has been shown to exhibit decreasing returns to scale due to difficulties in transportation (Baker, Chon et al. 2008). The two sub-technologies considered are *Gasification of biomass* and *High Efficiency Steam Cycles* (Baker, Chon et al. 2008). The Bio-Electricity technology has the advantage that it can be integrated with CCS to result in net carbon removal.

---

\(^3\) We define R&D endpoints as specific energy costs, efficiency and performance targets for each of the energy technologies.
**Liquid Bio-Fuels:** The advanced second generation Bio-Fuel technology involves the production of liquid fuel (ethanol, diesel, or gasoline) from non-food cellulosic biomass feedstock. The bio-mass to bio-fuels conversion process usually consists of two stages: the bio-mass feedstock is first broken down into a simpler by-product and then converted into liquid fuel. Seven processes were originally defined in Baker and Keisler (2011), including Stage 1 processes: *Selective thermal processing, acid and enzymatic hydrolysis, gasification* and Stage 2 processes: *refining, fermentation, aqueous phase processing* and *synthetic gas processing*. In order to simplify our analysis we reclassify the sub-technologies into four independent sub-technology paths, each comprised of a stage 1 and stage 2 process (*Selective Thermal Processing 1, Selective Thermal Processing 2, Gasification and Hydrolysis*). The details of these are shown in Olaleye (In Preparation).

**Batteries for Electric Transportation:** The sub-technologies considered consist of the *Lithium Ion* and *Lithium Metal* batteries (Baker, Chon et al. 2010).

We rely on probability elicitations from Baker et al. (Baker, Chon et al. 2008, Baker, Chon et al. 2008, Baker, Chon et al. 2009, Baker, Chon et al. 2009, Baker, Chon et al. 2010, Baker and Keisler 2011) to determine the probability of technological breakthroughs conditional on R&D investment. For each set of technologies a target endpoint was defined, and the probability of achieving this target was elicited from energy experts with care taken to adjust for human decision making biases by various means, including using a number of experts, a diverse pool of experts and an appropriate aggregation method (Baker and Olaleye 2013). See (Baker and Solak 2011, Baker and Olaleye 2013) for a reference for the defined funding levels for the CCS, Nuclear and
Solar technologies. The results of the elicitations are summarized in the appendix in Table A.1.

3.3 Methodology

In this section we outline the problem formulation, the models used and their calibration. We also detail the approach for the regression study and the methodology for the probabilistic scenario analysis conducted over a set of R&D portfolios. We define an energy scenario as a specific set of cost and performance parameter values for each of the 6 energy technologies considered. Using endpoints defined in the expert elicitations, these 6 technologies lead to a total of 3780 energy scenarios. We define an R&D portfolio as a specific set of funding levels for each of the 6 technologies. Each portfolio is associated with a probability distribution over the energy scenarios.

As earlier discussed, the goals of this paper include the following: to understand the relative importance of individual energy technology advancements in an economy facing uncertain climate damages; to examine the interactions between pairs of advanced technologies and how these interactions change with uncertainty in climate damages; and to examine for stochastic dominance between R&D portfolios by examining the cumulative distribution function (CDF) of these R&D portfolios over the energy scenarios. To achieve these goals, we do the following: we use a modified DICE model to conduct a global scenario analysis by examining all the possible energy scenarios that can result from our mix of technologies; based on the result of the scenario analysis, we conduct a regression analysis to examine interactions between technology pairs under damage uncertainty; and we finally conduct a stochastic dominance analysis to evaluate how CDF’s of some R&D portfolios change given the different damage uncertainty
cases. The rest of the methodology section is organized as follows: In section 3.3.1, we discuss the problem formulation of the stochastically reformulated DICE model, the implementation of technological change in GCAM, the parameterization and integration of technological change into the DICE model and the modeling of the uncertain climate states. In order to examine the nature of the interaction effects between energy technologies, we also conduct a regression study, described in section 3.3.2.

3.3.1 Problem Formulation

We use a two-stage decoupled model of uncertainty and learning. We discuss merits and limitations of our approach in the literature review and conclusion sections. We evaluate each of the 3780 future energy scenarios that can result from our portfolio of technologies in a modified version of the DICE model that includes uncertainty and learning about climate damages. The original DICE model, developed by William Nordhaus, is a top down neoclassical economic growth model of the climate and the economy (Nordhaus 2008). It has a single global region, with population and labor productivity modeled as exogenous parameters. It solves by optimizing the flow of consumption discounted by a rate of time preference (Nordhaus 2007, Nordhaus 2008).

The problem formulation for the stochastic reformulated DICE model that captures damage uncertainty is given in Equations (1) to (10) and a description of the variables and parameters used is given in Table 1 (exponents are given in parenthesis to differentiate from superscripts). This formulation builds on the formulation by (Baker and Solak 2013). Equations (1)-(8) are identical to the original DICE model, except that the variables are conditioned by the climate damage state of the world, \( \omega \). Equations (9) and (10) are non-anticipativity constraints. Technical advancement from R&D induced
technical change is modeled to kick in at 2055, and similarly, uncertainty about the severity of climate damage is assumed to be resolved by 2055. That is, decisions are made under uncertainty about climate damages during the 2005–2055 stage; the levels of the climate damages are explicitly known from 2055.

$$\max_{\mu^o, x^o} \sum_{t=0}^{5} p^o \sum_{t} R_t u_t^o$$

$$u_t^o = L_t^o \left( \frac{\theta_t^o}{\theta_t^o} \right)^{(1-\beta)} - 1 \quad \forall t, \omega$$

$$y_t^o = \theta_t^o + l_t^o \quad \forall t, \omega$$

$$k_t^o = l_{t-1}^o + (1-\sigma)k_{t-1}^o \quad \forall t, \omega$$

$$\tau_t^o = H\left(\tau_{t-1}^o, e_t^o\right) \quad \forall t, \omega$$

$$y_{t}^{e_{o}} = A_l^o \theta_t^o (1-\gamma) \theta_t^o (r_{t})^o \quad \forall t, \omega$$

$$y_{t}^{e_{o}} = \frac{1-\left(c_{\theta}^\dagger (\mu_t^o)\right)}{D_r^o (\tau_t^o)} y_t^{e_{o}} \quad \forall t, \omega$$

$$e_t^o = S_t^o (1 - \mu_t^o) y_t^{e_{o}} + E_t^o \quad \forall t, \omega$$

$$k_t^o - \sum_{\omega \in \Omega} p^o k_t^o = 0 \quad \forall t \leq 5, \omega$$

$$u_t^o - \sum_{\omega \in \Omega} p^o u_t^o = 0 \quad \forall t \leq 5, \omega$$
Table 1: Parameters and variables of the scenario evaluation model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_t$</td>
<td>$x$</td>
</tr>
<tr>
<td>Utility discount factor for period t</td>
<td>All other decision variables aside $\mu$</td>
</tr>
<tr>
<td>$A_t$</td>
<td>$o_t$</td>
</tr>
<tr>
<td>Level of total factor productivity in period t</td>
<td>Consumption in period t</td>
</tr>
<tr>
<td>$S_t$</td>
<td>$y_t$</td>
</tr>
<tr>
<td>Ratio of emissions to output in period t</td>
<td>Net output of goods/services in period t</td>
</tr>
<tr>
<td>$E_t$</td>
<td>$y^g_t$</td>
</tr>
<tr>
<td>Emissions from deforestation in period t</td>
<td>Unadjusted output in period t</td>
</tr>
<tr>
<td>$L_t$</td>
<td>$k_t$</td>
</tr>
<tr>
<td>Population and labor input in period t</td>
<td>Capital stock in period t</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$l_t$</td>
</tr>
<tr>
<td>Elasticity of marginal utility of consumption</td>
<td>Investment in period t</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$e_t$</td>
</tr>
<tr>
<td>Elasticity of output with respect to capital</td>
<td>Total carbon emissions in period t</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$\mu_t$</td>
</tr>
<tr>
<td>Rate of depreciation of capital</td>
<td>Emissions abatement in period t</td>
</tr>
<tr>
<td>$u_t$</td>
<td>$\tau_t$</td>
</tr>
<tr>
<td>Social utility in period t</td>
<td>Atmospheric temperature in period t</td>
</tr>
<tr>
<td>$\alpha_{0-3}$</td>
<td>$\zeta$</td>
</tr>
<tr>
<td>Marginal Abatement Cost (MAC) Calibration parameters</td>
<td>Energy scenario</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>$\omega$</td>
</tr>
<tr>
<td>Collection of all climate damage states</td>
<td>Climate damages state. $\omega, \omega' \in \Omega$</td>
</tr>
<tr>
<td>$p^\omega$</td>
<td>Probability of damage state $\omega$</td>
</tr>
</tbody>
</table>

The objective of our model is to maximize the expected value of societal utility over all states of the world, denoted $\omega \in \Omega$, where the states of the world reflect the different possible climate damages. Equation (1) gives the objective function which maximizes the discounted sum of per capita utility of consumption. Equations (2) – (8) represent equations of the original DICE model (Nordhaus 2008), with the modification that the variables of each constraint are conditioned by the climate damage state. The per-period utility is defined in Equation (2) as an isoelastic function of the rate of consumption with $\beta$ the constant elasticity of the marginal utility of consumption. Equations (3) and (4) give the consumption and capital balance equations respectively.
Equation (5) connects climate emissions and economic activity recursively to the carbon cycle and physical system. The function $H(\bullet)$ represents the relationship between future atmospheric temperature and both the current atmospheric temperature and current emission trajectory. The gross economic output, given by Equation (6), is a Cobb-Douglas function of the capital stock $k_t$ and labor force $L_t$, augmented by productivity $A_t$. Equation (7) represents the effect of climate damages and the cost of carbon abatement on the gross economic output $y_t$ to give the net economic output. As the damage function $D_D(\varepsilon_t^{\omega})$ increases with temperature rise and the abatement cost function $c_D^*(\mu_t^{\omega})$ increases with increased abatement, unadjusted output reduces.

Equation (8) links the carbon emissions $e_t$ and the abatement control rate $\mu_t$. As we have formulated this model as a stochastic programming problem, the first stage variables are constrained by non-anticipativity constraints in equations (9) and (10). Note that we sum over $\omega'$ in $\Omega$ to differentiate between the variables in the first and second stages. These ensure that decisions taken in the first stage (pre 2055) are the same for all the realizations of the climate damages state. Baker & Solak (2013) show that these two constraints are sufficient to enforce non-anticipativity in the model.

### 3.3.1.1 Implementing technical change.

In this section, we discuss the incorporation of technological advancement into the DICE model. We model technical change as impacting the cost of abatement starting from the year 2055. The reference abatement cost function $c_D(\mu_t)$ in the original DICE model is given in Equation (11), with $\tau_t$ representing the regional participation cost ratio.
and $\theta_1, \theta_2$ calibrated parameters of the abatement cost function. We modify the cost function by estimating the impact of technical advancement on the marginal abatement cost (MAC) curve for each scenario $\zeta$, where the MAC is the cost of abating one extra ton of carbon emissions. The MAC derived from Equation (11) is shown in (12). Each technology scenario $\zeta$ is associated with a set of parameters $\alpha_{0:3}^\zeta$ through $\alpha_3^\zeta$, which impact the MAC multiplicatively through the expression

$$\left[1-(\alpha_0^\zeta + \alpha_1^\zeta \mu_t + \alpha_2^\zeta \mu_t^2 + \alpha_3^\zeta \mu_t^3)\right].$$

The resulting cost function is derived by integrating the MAC. It is given in Equation (13) and its derivation is explicitly shown in Appendix B. The calibration of the $\alpha_0^\zeta$ .... $\alpha_3^\zeta$ parameters, which are estimated using the GCAM model, is discussed in section 3.3.1.2.

$$c_D(\mu_t) = \pi_t \theta_{t1} \mu_t^{\theta_1}$$

(11)

$$MAC(\mu_t) = \pi_t \theta_{t1} \theta_2 \mu_t^{\theta_1 - 1}$$

(12)

$$c_D^\zeta(\mu_t) = \pi_t \theta_{t1} \theta_2 \mu_t^{\theta_1} \left[\frac{1-\alpha_0^\zeta - \mu_t \alpha_1^\zeta}{1 + \theta_2} - \frac{(\mu_t)^2 \alpha_2^\zeta}{2 + \theta_2} - \frac{(\mu_t)^3 \alpha_3^\zeta}{3 + \theta_2}\right] \forall t, \omega$$

(13)

### 3.3.1.2 Parameterizing technical change using the Global Change Assessment Model

In order to assess the impact of technological advancement on the MAC and estimate the $\alpha_0^\zeta$ .... $\alpha_3^\zeta$ parameters, we use the GCAM model. Here, we briefly discuss the GCAM model, the generation of the MACs for the energy scenarios, and the parameterization of the resulting MAC curves.

The GCAM model is a large integrated assessment energy-economic model composed of interacting agricultural, climate, land use and economic units (Edmonds,
Wise et al. 1994, Clarke, Kyle et al. 2008). The model, developed by the Joint Global
Change Research Institute of the Pacific Northwest Laboratory in affiliation with the
University of Maryland, has been used in several studies including the 4th and 5th
Intergovernmental Panel on Climate Change (IPCC) assessment reports. It is
technologically detailed and hence readily amends itself to modeling different future
energy scenarios.

To derive the MAC curves that serve as input to the DICE model, Zdybel (2013)
modelled the energy technologies in GCAM, based on the same cost and performance
parameters used in this study (Baker, Chon et al. 2008, Baker and Keisler 2011, Baker,
2010). The technologies are assumed to diffuse globally. The MACs are derived from
GCAM in the method described in (Weyant and Hill 1999). Note that while GCAM has
5-year time steps, DICE has 10-year time steps, starting in 2005. Zdybel (2013) modeled
the energy technology scenarios with respect to 2050; we integrate these MAC’s into
DICE at 2055. The default MAC’s are used in DICE before 2055; thereafter we use the
R&D induced technologically adjusted MACs. Some energy scenarios MACs are shown
in Figure 2.

**MAC Parameterization:** Zdybel (2013) parameterized each of the resulting
MACs, in order to facilitate the use of the MACs as inputs to the DICE model. The
parameterization is conducted using a polynomial, order three, regression estimation
technique (Zdybel 2013), leading to estimated values for the parameters \( \alpha_0^c, \alpha_1^c, \alpha_2^c, \alpha_3^c \)
used in equation (13) above.
Figure 2: Marginal abatement cost measured as percentage of gross domestic product GDP.

Curves of some selected energy scenarios generated by GCAM for the year 2050 and integrated into DICE at 2055 (shown as a percentage of 2055 GDP values). It shows the effect of technological advancement (Adv.) in selected individual technologies. Biofuels, Batteries for Transportation and Bio-Electricity MAC’s are similar to Solar’s and are excluded for clarity.

3.3.1.3 Uncertainty in climate damages.

A major source of uncertainty in climate change is the severity of resulting climate damages. To understand how uncertainty in climate damages affects the optimal near term responses, Baker & Solak (2011, 2013) model climate damage uncertainty using mean preserving spreads. The base, No-Uncertainty case is taken from (Nordhaus 2008) and assumes a 1.1% loss in GDP given a 2°C rise in mean atmospheric temperature. We consider Medium and High Uncertainty cases, described in Table 2 with \( \Pi_2 \) denoting the calibrated parameter from the damage function given by

\[
D_p(\tau^*) = [1 + \Pi_1 \tau^* + \Pi_2 \tau^*] \quad \text{and} \quad \Pi_1 \text{ set to 0.}
\]

We implement uncertainty with learning: near
term decisions are made under uncertainty; later decisions are made based on the revealed state of the world in 2055. To examine the effect of early or late resolution of damage uncertainty, we also examine the effect of climate damage resolution in 2035 and 2075 relative to 2055.

Table 2: Damage Uncertainty Cases.

The table shows the percentage GDP loss for a 2°C rise in mean atmospheric temperature for each of the uncertainty cases. The Medium and High Uncertainty cases are modeled as mean preserving spreads of the No Uncertainty case with the appropriate probabilities shown in the second row. Π denotes the calibrated parameter from the damage equation that corresponds to each of the GDP losses.

<table>
<thead>
<tr>
<th></th>
<th>No Uncertainty</th>
<th>Medium Uncertainty</th>
<th>High Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP Loss (%)</td>
<td>1.1</td>
<td>0</td>
<td>3.3</td>
</tr>
<tr>
<td>Probability (%)</td>
<td>100</td>
<td>66.7</td>
<td>33.3</td>
</tr>
<tr>
<td>Π</td>
<td>0.003</td>
<td>0</td>
<td>0.009</td>
</tr>
</tbody>
</table>

3.3.2 Regression Methodology

We conduct a regression analysis in order to examine the relationships between technologies. We evaluate the contribution of each energy technology and each technology pair to the change in the societal utility across the 3780 energy scenarios. We limit the regression to the second order as third and higher order interactions are statistically insignificant, possibly due to the number of observations used. The formulation for this study is given in Equation (14), where \( \alpha \) is the intercept coefficient, \( \alpha_i \) and \( \alpha_{ij} \) represents the coefficients indicating the magnitude of the contribution due to technological advancement of the different energy technologies and energy technology pairs to the societal utility respectively, and \( \Delta U \) is the change in the societal utility with respect to the ‘no technical advancement’ scenario. The independent variables are index
variables representing technological advancement in the technologies and technology pairs. As an example, $\xi_{CCS}$ is an index variable indicating the level of technological development in the CCS technology, with 0 for No technological advancement and 1 for the best CCS sub-technological success level Post-Combustion based on its cost. The data for the No technological advancement costs is from GCAM version 3.0 from Clarke et al. (Clarke, Lurz et al. 2007); the remainder of the data is from the elicitation papers (Baker, Chon et al. 2008, Baker and Keisler 2011, Baker, Chon et al. 2009, Baker, Chon et al. 2008, Baker, Chon et al. 2009, Baker, Chon et al. 2010) as summarized in Zdybel (Zdybel 2013). These data are summarized in the appendix in the last column of Table A.1.

The pair variables, such as $\xi_{CCS:NUC}$ are simply the product of the two indicator variables, in order to get at the interaction of the two.

$$
\Delta U = \left( \alpha_0 + \alpha_1 \xi_{CCS} + \alpha_2 \xi_{NUC} + \alpha_3 \xi_{BAT} + \alpha_4 \xi_{BF} + \alpha_5 \xi_{BE} + \alpha_6 \xi_{SOL} + \alpha_7 \xi_{CCS:NUC} + \alpha_8 \xi_{CCS:BAT} \\
+ \alpha_9 \xi_{CCS:BF} + \alpha_{10} \xi_{CCS:BE} + \alpha_{11} \xi_{CCS:SOL} + \alpha_{12} \xi_{NUC:BAT} + \alpha_{13} \xi_{NUC:BF} + \alpha_{14} \xi_{NUC:BE} + \alpha_{15} \\
\xi_{NUC:SOL} + \alpha_{16} \xi_{BAT:BF} + \alpha_{17} \xi_{BAT:BE} + \alpha_{18} \xi_{BAT:SOL} + \alpha_{19} \xi_{BF:BE} + \alpha_{20} \xi_{BF:SOL} + \alpha_{21} \xi_{BE:SOL} \right) 
$$

(14)

We note that we use the regression analysis to summarize the hidden effects of the IAM model by estimating the relationship of the input data to the IAM model output utility. While this is not a conventional use of regression analysis, we argue that it is a valid methodology given that all inputs satisfy the underlying assumptions for a regression and the IAM model process implies that there is a causal relationship between the independent variables and dependent variable (Berry, 1993).

### 3.4 Key Results and Analysis

#### 3.4.1 Comprehensive Scenario Analysis

In this subsection, we present the results from the scenario analysis of all the energy
scenarios.

Figure 3 shows the expected utility of each of the energy scenarios for each of the three uncertainty cases, where the scenarios are presented in increasing order of societal utility for the No Uncertainty case: scenario 0 represents no technological development and scenario 3780 corresponds to the maximum possible development in each technology. Each scenario in the Medium and High Uncertainty is presented in the same ordering as in the No Uncertainty case. The societal utilities have been normalized to be the additional utility above the No Uncertainty, No technological advancement scenario.

From Table 3 we observe that the High and Medium Uncertainty scenarios have two separate horizontal bands. The lower bands of scenarios are scenarios without technological advancement in CCS (but with advancement in Nuclear). Therefore we find that energy scenarios lacking technological advancement in CCS perform more poorly in the highly uncertain damage cases compared to the No Uncertainty case. It appears that CCS serves as a hedge against climate damage uncertainty. This may be related to the fact that CCS has a significantly lower marginal abatement cost at full emissions abatement, as can be seen in Table 3.

The significance of CCS can also be noticed from the least developed scenarios (0-300), where we observe a substantial divide, especially in the Medium and High Uncertainty cases, between scenarios that do not have technological advancement in either CCS or Nuclear and all the other scenarios.
Figure 3: Expected societal utility of energy scenarios ordered by increasing utility in the No Uncertainty case.

The societal utility of the scenarios is referenced to the No Uncertainty No technological advancement scenario. Utility here is a measure of societal welfare and is in utility units (Nordhaus W., 2008) not GDP $ per person value.

In addition, consistent with the findings in Baker & Solak (2011), we note that the slope is greatest for the Medium Uncertainty case and lowest for the High Uncertainty case, implying that technological change has the most impact in the Medium Uncertainty case and the least in the High Uncertainty case. This is driven by two factors. First, the probability of high damages in the High Uncertainty case is quite low; technical change only has an impact when damages are high (as opposed to zero); therefore the impact of technical change is attenuated. Second, in the High Uncertainty case there is a cost-side benefit (better technologies make it less expensive to reduce emissions), but no
environmental-side benefit (the large damages mean that abatement is 100% regardless of the technology – better technology does not change the level of abatement or the environmental outcomes). In the Medium Uncertainty case, on the other hand, the probability of damages is larger, and both the cost-side and environmental-side benefits are large. See (Baker and Solak 2011) for further discussion.

3.4.2 The Timing of Uncertainty Resolution

As noted, we use a two stage model, with uncertainty resolved in 2055. Here we examine the effect of earlier or later resolution of climate damage uncertainty.

Figure 4 shows the results in the Medium Uncertainty case, under the assumptions that uncertainty is resolved in 2035, 2055 or 2075. The timing of the resolution of climate uncertainty not only affects the utility, as expected, but also has an effect on the relative value of technological advancement in the technologies. Scenarios with advancement in CCS increase in relative value with later resolution of uncertainty. This implies that the less confidence we have about understanding the full impact of climate change, the bigger the role of CCS.
Figure 4: Effect of early or late resolution of climate damages.

Scenarios here are ordered by 2055 expected utility in an economy facing Medium Uncertainty when climate damages are resolved at 2035, 2055 and 2075.

3.4.3 Regression Analysis for Substitution and Complementarity Effects Between Technologies

Table 3 and Figure 5 show the results of the regression study discussed in Section 3.3.2. Specifically, Table 3 shows the contribution of each technology, and of each statistically significant technology pair, to the societal utility under the three damage cases considered. The table also shows the $R^2$ and the residual error values for the three uncertainty cases. We note that these values show that a multiple linear regression is a relatively good fit for the model even though the multi-model framework (GCAM and DICE) is highly non-linear. The goodness of fit as predicted by the $R^2$ reduces in the Medium and High Uncertainty cases due to the effect of damage uncertainty on the utility of the different R&D endpoints.

Here we discuss the key results related to the regression study.
Table 3: Regression results for all uncertainty cases and statistically significant technology pairs.

Gray shaded cells represent the significantly complementary CCS:BE pair. * represents the degree of statistical significance of the results; *** $\Rightarrow p<0.001$, ** $\Rightarrow p<0.01$, * $\Rightarrow p<0.05$ and . $\Rightarrow p<0.1$. DF is the degree of freedom.

<table>
<thead>
<tr>
<th>Technology Pairs</th>
<th>No Uncertainty</th>
<th>Medium Uncertainty</th>
<th>High Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>18.6***</td>
<td>39.9***</td>
<td>58.1***</td>
</tr>
<tr>
<td>Technologies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCS</td>
<td>61.8***</td>
<td>115.0***</td>
<td>66.4***</td>
</tr>
<tr>
<td>NUC</td>
<td>128.0***</td>
<td>145.3***</td>
<td>46.3***</td>
</tr>
<tr>
<td>BAT</td>
<td>5.1</td>
<td>7.8</td>
<td>5.7</td>
</tr>
<tr>
<td>BF</td>
<td>9.2*</td>
<td>16.8*</td>
<td>7.4*</td>
</tr>
<tr>
<td>BE</td>
<td>22.3***</td>
<td>30.8***</td>
<td>6.0*</td>
</tr>
<tr>
<td>SOL</td>
<td>13.9***</td>
<td>7.8</td>
<td>-0.2</td>
</tr>
<tr>
<td>Technology Pairs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCS:NUC</td>
<td>-7.1**</td>
<td>-32.8***</td>
<td>-25.1***</td>
</tr>
<tr>
<td>CCS:BE</td>
<td>33.7***</td>
<td>42.7***</td>
<td>10.6***</td>
</tr>
<tr>
<td>NUC:BE</td>
<td>-13.8***</td>
<td>-20.5***</td>
<td>-4.3*</td>
</tr>
<tr>
<td>NUC:SOL</td>
<td>-9.9***</td>
<td>-7.6</td>
<td>-0.5</td>
</tr>
<tr>
<td>BF:BE</td>
<td>-6.2*</td>
<td>-12.2*</td>
<td>-3.3</td>
</tr>
<tr>
<td>$R^2$</td>
<td>84%</td>
<td>69%</td>
<td>57%</td>
</tr>
<tr>
<td>Residual Error (DF=3758)</td>
<td>25</td>
<td>47</td>
<td>22</td>
</tr>
</tbody>
</table>
Figure 5: Regression results for statistically significant technologies and technology pairs.

From Figure 5 we see that CCS and Nuclear technologies have the largest impact on societal utility, though we should note that Nuclear technologies require a much higher level of R&D investment (See Table 4). As an example the maximum R&D funding requirement in the Nuclear technology is $15.4 billion, 30 times more than the maximum $519 million funding in CCS. Batteries for Electric Transportation have almost no impact and weak significance at the levels of technological advancement considered here.

The importance of technologies depends on the uncertainty in climate damages: Nuclear has the largest contribution to utility in both the No Uncertainty and Medium Uncertainty cases, but CCS has the largest contribution in the High Uncertainty case. On the other hand, Solar PV plays a role when there is no damage uncertainty, but has very little effect.
in Medium and High Uncertainty cases. The reason for these results can be seen by looking at the shape of the MACs that result from R&D success in solar PV or CCS (Figure 6). From Figure 6 (and Figure 2) we see that CCS pivots the MAC down: the marginal cost is relatively lower at high abatement levels, in contrast to Nuclear, which maintains a more of a constant reduction in the MAC. This means that when climate damages turn out to be very high, CCS has a large payoff. Solar is the opposite, with a relatively large impact at lower abatement levels: its MAC is below Bio-Fuels and Batteries for Electric Transportation at low levels of abatement, but above for higher abatement levels.

![Figure 6: Close-up of the abatement cost curves of Solar, Batteries for Electric Transportation and Bio-fuels technologies for the year 2050. These are shown in % 2055 GDP values.](image)

A positive interaction coefficient for a technology pair implies a complementary effect between the pair (Aiken & West, 1991): the higher the level of technological
advancement in one technology, the greater the utility of having technological advancement in the other technology that consists the technology pair. A negative interaction coefficient on the other hand implies substitutability across technologies, as the higher the level of technological advancement in one technology, the lower the additional utility of having increased technological advancement in the other technology. The CCS:Bio-Electricity pair are significant complements while most of the other technology pairs are substitutes: it is the only pair with significant positive coefficients. These two technologies are complements because together they lead to a path that can include negative emissions: the carbon captured during the lifetime of the bio-mass feedstock can be sequestered during the energy generation process with CCS. Most of the technologies are substitutes, however, as they are competing supply technologies (McJeon, Clarke et al. 2011). Nuclear energy is a substitute for CCS, Bio-Electricity and Solar energy technologies, while Bio-Fuels also substitutes Bio-Electricity as they share and compete for a common biomass feedstock (Luckow, Wise, Dooley, & Kim, 2010). We note that Nuclear:Bio-Electricity are significant substitutes in the No Uncertainty case with their degree of substitution reducing as climate damage uncertainty increases in contrast with the CCS:Nuclear pair whose degree of substitution increases with climate damage uncertainty. This is due to the MAC curves of the technologies with Nuclear resulting in a significant constant reduction from the reference MAC (as seen in Figure 6) while CCS leads to significant reductions in the MAC towards full carbon abatement. However, technology pair relationships do not qualitatively change as climate damage uncertainty increases.
Table 4: An illustration of the R&D effectiveness of the sub-technologies.

*Assessed at the medium R&D funding level, the sub-technology high success endpoint (defined in Table A.1) and the No Damage Uncertainty Case.*

<table>
<thead>
<tr>
<th>Technology</th>
<th>Sub Technology</th>
<th>Weighted Funding [Funding / Success Probability] ($million)</th>
<th>Utility</th>
<th>R&amp;D Effectiveness [Utility / Weighted Funding]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon Capture &amp; Storage</td>
<td>Post-Combustion</td>
<td>262</td>
<td>53</td>
<td>0.2016</td>
</tr>
<tr>
<td>Solar</td>
<td>In-Organic</td>
<td>266</td>
<td>13</td>
<td>0.0505</td>
</tr>
<tr>
<td>Carbon Capture &amp; Storage</td>
<td>Chemical Looping</td>
<td>280</td>
<td>76</td>
<td>0.2732</td>
</tr>
<tr>
<td>Bio-Electricity</td>
<td>Steam</td>
<td>414</td>
<td>5</td>
<td>0.0129</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Hydrolysis</td>
<td>427</td>
<td>2</td>
<td>0.0037</td>
</tr>
<tr>
<td>Nuclear</td>
<td>Light Water Reactor</td>
<td>769</td>
<td>122</td>
<td>0.1590</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Gasification</td>
<td>1,242</td>
<td>7</td>
<td>0.0055</td>
</tr>
<tr>
<td>Bio-Electricity</td>
<td>Gasification</td>
<td>1,268</td>
<td>9</td>
<td>0.0073</td>
</tr>
<tr>
<td>Carbon Capture &amp; Storage</td>
<td>Pre-Combustion</td>
<td>1,364</td>
<td>65</td>
<td>0.0478</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Sel. Thermal Proc. 1</td>
<td>1,514</td>
<td>2</td>
<td>0.0015</td>
</tr>
<tr>
<td>Batteries for Elect. Trans.</td>
<td>Lithium-Ion</td>
<td>1,691</td>
<td>4</td>
<td>0.0023</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Sel. Thermal Proc. 2</td>
<td>2,364</td>
<td>1</td>
<td>0.0003</td>
</tr>
<tr>
<td>Batteries for Elect. Trans.</td>
<td>Lithium-Metal</td>
<td>3,472</td>
<td>5</td>
<td>0.0014</td>
</tr>
<tr>
<td>Nuclear</td>
<td>High Temp Reactor</td>
<td>11,209</td>
<td>122</td>
<td>0.0109</td>
</tr>
<tr>
<td>Solar</td>
<td>Organic</td>
<td>24,692</td>
<td>13</td>
<td>0.0005</td>
</tr>
<tr>
<td>Nuclear</td>
<td>Fast Reactors</td>
<td>926,600</td>
<td>122</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Figure 7: Illustration of the R&D effectiveness of the sub-technologies.

The sub-technologies are assessed at medium funding with the probability of resulting in a high success endpoint (defined in Table A.1) under the No Damage Uncertainty Case.

To put these results in context, we provide Table 4 and Figure 7, which provides an illustration of the R&D effectiveness of the sub-technologies when they are funded at the medium funding level. The third column of the table shows what we are calling the weighted funding; this is the R&D funding amount divided by the probability of success. We note that in some cases R&D can result in two possible endpoints as shown in Table A.1; in this table we consider only the high success endpoint. While it is possible, as seen in Table 4, for R&D in different sub-technologies to result in the same endpoint (Table A.1), the probability of success will be different. The weighted funding is a measure of the funding per unit of probability; we have presented the sub-technologies in this order.
The fourth column shows the utility of the sub-technologies; this is the utility in DICE when only this sub-technology is successful and all the others fail. The fifth column presents a measure of the R&D effectiveness of the sub-technologies. It presents the ratio of the utility of the sub-technologies to their weighted funding, giving a measure of expected utility per dollar R&D funding for each sub-technology. This is also shown in percentage terms in Figure 7, we see that though certain technologies may result in high utility to the society given success, their attractiveness from an R&D investment perspective might be very low due to their R&D funding cost and the success probability. We see that though success in Nuclear leads to the highest utility, it also has a very high R&D investment requirement, leading to a low R&D effectiveness ratio. R&D in CCS technologies, however, provides the highest R&D effectiveness return on utility per R&D funding investment.

3.4.4 Probabilistic Scenarios Analysis of Selected Future Energy States

In this section we present results on the probability distributions over scenarios conditional on different R&D portfolios. We examine the CDF’s of different technology R&D portfolios, and examine them for stochastic dominance relations. Stochastic dominance is a form of stochastic preference ordering for ranking random variables, such as portfolios. We consider two types of stochastic dominance. A random variable $X$ is said to first order stochastically dominate (FOSD) $Y$, if all decision makers who prefer more to less, prefer $X$ to $Y$. We illustrate this visually in Figures 8a, which shows the cumulative distribution functions of some selected technology portfolios under the assumption of No damage Uncertainty, with thicker lines representing higher R&D funding. If one portfolio FOSD another, then its CDF will be entirely to the right. For
example, the thick blue dashed line ‘portfolio 1’ which represents the portfolio with a high investment in all technologies, FOSD all other portfolios.

We also consider second order stochastic dominance, which is related to risk aversion. A random variable X second order stochastically dominates (SOSD) Y if all risk averse decision makers prefer X to Y. It can be visualized in Figures 8a as follows: if the area below the CDF for X measured from left to right, is less than the area below the CDF for Y measured from left to right, X SOSD Y. As an example, the red dashed line portfolio 2, with a funding amount of 27.5 billion, SOSD the dashed blue line representing portfolio 4 with a funding amount of 9.7, but does not FOSD it.

We highlight the dominance characteristics of these technology portfolios in Table 5. These portfolios are selected to get a view of the diverse nature of the sample space: we take into consideration our findings in this paper on complementarities and substitutes, as well as policy concerns such as focusing on renewables, funding constraints, and possible socio-political constraints on certain technologies. Portfolios 2, 9 and 11 look at the role of CCS and bioelectricity; portfolios 3, 4, and 7 look at the role of renewables versus non-renewables; portfolios 1,6,8, and 10 look at different funding levels; and portfolio 5 limits nuclear and excludes biofuels. In the table, we rank the considered portfolios in descending order of R&D funding and discuss properties of the CDF’s of the different R&D portfolios.

**Importance of directed R&D funding**: From Figures 8a, we observe that a large R&D Portfolio, portfolio 2, with a funding amount of 27.5 billion, only first order stochastically dominates one of the 11 selected portfolios despite having over 93% of the maximum possible R&D funding. It is dominated by most of the other portfolios: first
order stochastically dominated by four and second order stochastically dominated by six portfolios. We see in particular that even R&D portfolio 8 with a funding amount of 2.0 billion first order stochastically dominates it. This is likely because portfolio 2 does not include CCS or Bio-Electricity, the complementary pair. This shows that there are some very inefficient portfolios in the choice set, with the presence of funding in certain technologies a critical factor in the performance of these portfolios.

Reinforcing this, we see that technology portfolios 9 and 11, which have high CCS investment and relatively low R&D funding, first order stochastically dominate higher cost portfolios 4 and 7, which have no investment in CCS. This indicates that technical advancement in CCS may be important.

**Diminishing returns to scale:** We find that a portfolio that limits the maximum available funding to each technology to $1 billion, portfolio 8, with a funding amount of 7.0 billion, first order stochastically dominates all but the three largest portfolios (portfolios 1, 2 and 3 with R&D funding 29.4, 27.5 and 19.8 billion respectively). It also has second order stochastic dominance over one of those three, and is only FOSD by the maximum funding portfolio. Thus, it appears that this is a very efficient portfolio, and there may be significant diminishing returns to scale.

**A renewable only-portfolio is not favorable:** A renewable only R&D portfolio, portfolio 7, with a funding amount of 7.5 billion, is first order stochastically dominated by 10 of the 11 portfolios considered and is second order stochastically dominated by all the portfolios considered, including four portfolios that have lower funding amounts.

**Role of R&D investment in CCS increases with damage uncertainty:** Figures 8b shows the CDF’s of the same set of R&D portfolios in the High Uncertainty case (the
Medium Uncertainty case is not shown as the dominance relationships are similar to the other two uncertainty cases). Note that the horizontal axis is quite different in this case – expected utility is much higher in the High Uncertainty case than the No Uncertainty case. This figure illustrates that there is a slightly different picture when damage uncertainty is considered. As climate damage uncertainty increases, R&D portfolios lacking investment in CCS move to the left relative to the other portfolios, becoming less desirable from a dominance viewpoint. Note that the portfolios that do not include CCS are dominated by even more portfolios when we consider the High Uncertainty case: portfolio 2, with a funding amount of 27.5 billion, is FOSD by six portfolios in the High Uncertainty case, compared to only two in the No Uncertainty case; and portfolio 7, with a funding amount of 7.5 billion, is now dominated by all.
Table 5: Characteristics of selected portfolios.

FOSD: First Order Stochastically Dominates, FOSD.: First Order Stochastically Dominated, SOSD: Second Order Stochastically Dominates, SOSD.: Second Order Stochastically Dominated. *: Solar and Batteries for Electric Transportation technologies are defined only at High and Medium funding. Color shades represent funding level.

<table>
<thead>
<tr>
<th>Fund. Rank/Port. No.</th>
<th>Portfolios</th>
<th>R&amp;D Fund ($Bil)</th>
<th>Dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HIGH</td>
<td>MEDIUM</td>
<td>LOW</td>
</tr>
<tr>
<td>1</td>
<td>All High</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>2</td>
<td>All but Bio-Electricity &amp; CCS</td>
<td>27.5</td>
<td>FOSD only #7 (FOSD. by four ports.)</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>3</td>
<td>No Renewable</td>
<td>19.8</td>
<td>FOSD four ports. (FOSD. only by #1)</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>4</td>
<td>Renewable with Low CCS &amp; NUC</td>
<td>9.7</td>
<td>(FOSD. by four ports.) (SOSD. by three ports.)</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>5</td>
<td>No Biomass and Limited NUC</td>
<td>9.5</td>
<td>FOSD three ports. (FOSD. only by #1 &amp; #8)</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>6</td>
<td>Medium</td>
<td>9.4</td>
<td>FOSD four ports. (FOSD. only by #1 &amp; #8)</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>7</td>
<td>Renewables ONLY</td>
<td>7.5</td>
<td>(FOSD. by all but #10) (SOSD. by all)</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>8</td>
<td>All Sub-Technology less than 1 Billion</td>
<td>7.0</td>
<td>FOSD all but three ports.</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>9</td>
<td>Bio-Electricity &amp; CCS with others</td>
<td>5.2</td>
<td>FOSD four ports. including #4 &amp; #7</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
<tr>
<td>10</td>
<td>Low</td>
<td>2.6</td>
<td>FOSD none</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT*</td>
</tr>
<tr>
<td>11</td>
<td>Bio-Electricity &amp; CCS only</td>
<td>1.9</td>
<td>FOSD #7</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>BF</td>
<td>BAT</td>
</tr>
</tbody>
</table>

R&D has more value in the High Uncertainty case: From comparing Figures 8a and b, we find that with an increase in uncertainty of climate damages there exist a
larger number of strict dominance relations between the R&D portfolios. We observe that in the High Uncertainty case, most of the portfolios considered clearly stochastically dominate others or are dominated; in comparison with the No Uncertainty case, where there is less strict dominance. This implies that R&D has more value as the uncertainty in climate damages increases and that the need for directed R&D spending becomes more important as uncertainty increases.
(8a) - No Climate Damage Uncertainty Case

(8b) - High Climate Damage Uncertainty Case

Figures 8a and 8b: CDF’s of some selected energy scenarios in the No Uncertainty (8a) and High Uncertainty (8b) cases.
The R&D funding amount required for each of the scenarios is given in the legend.

3.5 Conclusion

In this chapter of the thesis, we use the results of previous expert elicitations as a starting point for developing future technology scenarios. We run the scenarios through a detailed Integrated Assessment Model, GCAM, in order to estimate the impact that different future energy scenarios have on the cost of mitigating climate change. We then use these estimations to run a stochastic version of a top-down IAM, DICE, to estimate the impact that the different scenarios have on social welfare in a world with uncertainty and learning about climate damages.

Findings from our study can be classified into two broad categories, those that corroborate and extend findings from prior studies and relatively novel findings. Similar to findings in McJeon et al. (2011), our analysis highlights the role of CCS as a significant technology in the response to climate change due to its low cost carbon abatement properties. We confirm that this result holds from an economic perspective and go further to extend it, by showing that the role of CCS in the portfolio depends on when climate damages are known for certain. We find that there exist significant decreasing marginal returns to R&D investment in some technologies [e.g. (Cohen & Levinthal, 1989)]; we go further to show that there is significant value in carefully choosing a portfolio. We also confirm that the impact of technological change is non monotonic in climate damage uncertainty, as shown in Baker & Solak (2011), as technological change has the most impact in the Medium Uncertainty case followed by the High and No Uncertainty cases. We however show that climate damage uncertainty has an impact on the role of the different technologies.
Beyond these findings that extend insights from prior studies, we have a set of novel results. We find that the expected impact of energy technologies in the economy depends on how much uncertainty there is around climate damages, with CCS increasing in value as uncertainty increases, and Solar decreasing in value. Given that there is deep uncertainty about climate damages (Groves and Lempert 2007), this implies that it may make sense to diversify a portfolio around different uncertainty scenarios. In a similar vein, we show that Bio-Electricity and CCS are significant complements, implying that technology portfolios may want to include both of these technologies if climate damage uncertainty is high. On the other hand, most technology pairs are substitutes including Bio-Electricity and Nuclear, which are strong substitutes.

The probabilistic scenario analysis also resulted in a series of new insights. First, we illustrate that it is not enough to just throw money at this problem, since some very expensive portfolios are dominated by much less expensive portfolios. For example, the expensive portfolio with all renewables is dominated by a less expensive portfolio that contains only Bio-Electricity and CCS. We also find that certain portfolios perform relatively well irrespective of the distribution of climate damages (such as the portfolios with medium funding for all technologies or with a cap of $1 billion, which are robust to climate damage uncertainty), while other portfolios are best under a particular climate uncertainty case (such as the portfolios dominated by Bio-Electricity and CCS, which perform well in the High Climate Damage Uncertainty case). This analysis also shows that there are significant decreasing marginal returns to scale among the portfolios. For example, a limited funding portfolio – capped at $1 billion per sub-technology – provides excellent returns to R&D investment compared to the maximum funding portfolio.
While this study was based on a set of 6 technologies, the key insight – that there is significant value in carefully choosing a portfolio – is likely to hold given a wider range of technologies. Similarly, the relationship between specific pairs of technologies is not likely to depend heavily on other technologies. On the other hand, it is possible that specific technologies, especially CCS, may play a smaller role if a broader portfolio of technologies is included. Thus, key goals for future research would be to (1) gather expert judgments on a wider range of energy technologies; and (2) incorporate the technologies into portfolio analysis, with an emphasis on understanding how the introduction of new technologies is likely to change the optimal portfolio.

Similarly, the results in this paper are driven by the specific Integrated Assessment Models we use, namely GCAM and DICE. While both of these models have been highly influential in the assessment of climate change policies, they are subject to the limitations of all models. Thus, another avenue for future work is to test the degree to which the specific relationships found in this study (such as the complementarity of CCS and bio-electricity) will hold when analyzed by a range of models. See Bosetti, et al. (2014) for a study comparing the sensitivity of several IAMs outputs to different energy input cost assumptions.

In order to make the model computationally tractable, we decoupled and sequentially resolved the two different uncertainties (technological change and climate damages). We note that while this work represents a promising approach to characterizing uncertainties in IAMs [as discussed by e.g. Kann & Weyant (2000)], there is significant potential for further future research to explore the effect of such disaggregated approaches to treating uncertainties in IAMs.

54
Related to this, there is question of the timing of the uncertainty resolution stages in IAMs [e.g. (Keller, Bolker, & Bradford, 2004)]. We briefly investigated and found that later resolution does impact the relative value of some of the technologies, with CCS gaining value the later the resolution of uncertainty. Finally, we note that this study does not consider the cost of R&D; future work can evaluate the effect of integrating R&D costs into a multi-IAM framework.
CHAPTER 4

PART 2: R&D ALLOCATION IN A LARGE PORTFOLIO OF 6 CLEAN ENERGY TECHNOLOGIES

4.1 Introduction

In this chapter we approach the R&D portfolio allocation problem from an optimization viewpoint, this is in contrast to the previous chapter where we approached the R&D technology policy problem from a comprehensive scenario analysis viewpoint. Specifically in this chapter we aim to determine the allocation of R&D funds to the competing energy technologies that maximizes the expected utility to the society.

We consider a set of advanced “not business as usual” cost and technical improvement endpoints\(^4\) in Solar, Nuclear, Carbon capture and storage (CCS), Liquid Bio-Fuels, Electricity from Biomass and Batteries for transportation, thereby extending previous work by (Baker and Solak 2011) who consider a portfolio of three technologies CCS, Nuclear and Solar. The uncertain nature of the possible climate damages and the uncertainty of the R&D process make this form of energy R&D portfolio allocation a very large scale computation problem. The scale of the problem exponentially increases when additional technologies are considered. The details of the study are discussed below.

\(^4\) Where an endpoint is a cost, and efficiency improvement, target for a particular sub-technology
4.1.1 Approach

Technological R&D in the energy industry has led to more efficient and economical energy systems [e.g. (Joskow & Rose, 1985)], resulting in significant improvement in societal welfare. The challenge of climate change makes the need for such guided technological innovation paramount in developing efficient cost competitive low carbon energy technologies. Given the paucity of information about the future, robust near term policy must take into account the uncertainty surrounding the impact of climate change, the uncertainty of the technological research and development (R&D) process and various uncertainties about the energy system. This paper focuses on determining the optimal technology policy response, in terms of the allocation of R&D investment across competing clean energy technologies.

We consider a set of potential R&D-induced technological breakthroughs in each of six technologies: Solar Photovoltaic, Nuclear Fission, Carbon Capture and Storage (CCS), Liquid Bio-Fuels, Electricity from Biomass and Batteries for Electric Transportation. Within each of these energy technologies, we consider different R&D initiatives that can lead to different possible variants of the technology e.g. Solar energy consists of projects including; Organic solar cells, In-Organic solar cells and 3rd Generation. The nature of such R&D quests is such that they are inherently uncertain (Pindyck, 1991), as R&D investments do not always guarantee technological advancement. Also little is known about the severity of the eventual damages from climate change. Therefore the optimal energy policy has to consider these two uncertainties. Investment in R&D must be made now to combat future unknown damages. This is because R&D investments can lead to different possible future energy
states of the world depending on the realization of the R&D technological development process and the climate damages.

We note that while this chapter addresses this problem from a portfolio optimization perspective, the previous study chapter (Olaleye & Baker, 2015) approached the near term energy policy problem from a global scenario analysis and stochastic dominance perspective. Some of the significant results of the previous chapter include a characterization of certain technology pairs as substitutes or complements; that R&D portfolios with CCS stochastically dominate R&D portfolios without CCS; and that there is significant value to having R&D investment in CCS especially when the possibility of climate damages are significant. We note that due to the scale of the problem, while the previous study considered a global scenario analysis of the energy scenarios, it only considered a very limited analysis of a few interesting energy portfolios. This chapter also builds extensively on previous work by Baker and Solak (2011), who consider portfolio optimization over a portfolio of three technologies - CCS, Nuclear and Solar - in the face of uncertain climate damages. There therefore remains the question of the efficient allocation of R&D investments for this set of clean energy technologies, given an R&D budget.

Another question addressed in this chapter is finding the optimal solution to a sub-problem of the R&D portfolio allocation problem, when a solution to the original problem already exists. This is useful when there is a need to exclude a technology (or set of technologies) in the portfolio. This chapter provides tractable solutions to both of these questions.
Portfolio optimization problems are known to be NP-complete knapsack bin-packing problems (e.g. (Garey & Johnson, 1979)). Some methods exist for reducing the complexity of these knapsack problems, including dominance relation techniques (Andonov et al., 2000; Poirriez et al., 2009). We show in (Olaleye and Baker, 2015) that while such dominance relations exists between some energy portfolios in our 6 energy technology problem, such orderings may however not always hold across all the R&D portfolios, and that to exhaustively evaluate all possible portfolios is computationally more intensive than exhaustively solving the portfolio problem. We also note that our particular problem has the characteristic that the sub-technologies are not independent of each other, hence stochastic dominance techniques cannot be used across all portfolios. We therefore approach this problem from a genetic algorithm optimization perspective given the scale and the nature of the problem e.g Hassan, Cohanim et al. 2005). Our approach, discussed briefly in the introduction section, is detailed below.

Our approach consists of a multi-model sequential optimization framework composed of an R&D investment portfolio optimization model developed herein and two integrated assessment models; the Global Change Assessment Model (GCAM (Clarke et al., 2007) and a stochastic reformulation of the Dynamic Integrated Climate-Economic model (DICE (Nordhaus, 1993; Baker and Solak, 2014)).

First, technological R&D advancement is modeled deterministically in the GCAM model. This is done by modeling the future energy scenarios, resulting from combinations of the different advanced technologies in the technologically detailed GCAM model. Secondly, the economic impact of these energy technology scenarios are evaluated under different assumptions of climate damages impact in the stochastically
reformulated DICE model. Finally, using the R&D investment portfolio optimization model, we determine the optimal R&D funding allocations to the different technologies, with each specific R&D budget providing a conditional distribution over the likelihood of the different energy scenarios, based on prior expert elicitation studies from Baker et al. (2008b; 2008a; 2009a; 2009b; 2010; 2011). The large number of R&D projects considered coupled with the uncertainty in both the climate damages and the R&D process makes the portfolio allocation a relatively large scale computation problem. This restricts us to a sequential multi-model approach, as the problem is intractable in a single optimization model.

Additionally, given that there might exist a need to re-balance a portfolio by removing one of the technologies considered, we propose a methodology for solving sub problems of the portfolio problem, based on the solutions to the initial portfolio optimization problem and the interaction characteristics of the excluded technology. We approach this by decomposing the portfolio problem and showing that the optimality conditions are the same for both the original portfolio problem and the sub-problem, provided that the excluded technology does not have any interaction effect with any of the other technologies (or set of technologies) in the portfolio. This shows that there is no need to re-optimize a sub-portfolio provided no interactions exist with the excluded technology.

We discuss the problem scope and the specific approach for the R&D portfolio optimization problem in Section 4.2. In Section 4.3, we discuss the methodology and we also discuss our approach to finding solutions to subsets of the R&D portfolio problem (de-compartmentalization) in Section 4.3.4.
4.2 Problem Scope and Solution Approach

4.2.1 Problem Scope

Given a specific R&D investment budget, there exist different possible allocations of R&D funds across projects. We define an R&D portfolio as a specific set of funding levels for each of the technology projects considered. One of our goals is to determine efficient portfolios, which maximize the economic societal utility of technological advancement through R&D in the face of climate change. Depending on the R&D budget, and given our set of technologies, there can be over 2 billion possible R&D portfolios. Due to the uncertainty in the R&D process, thousands of different future energy scenarios can result from any of these specific R&D portfolios. An energy scenario is defined as a specific set of cost and technical performance specifications for each of the 6 energy technologies considered; given our data there exist 3,780 such scenarios in our study. Each R&D portfolio has a probability distribution over the set of energy scenarios. To model the uncertainty in climate damages, we consider three different distributions of the climate state: modeled as increasing risk (Rothschild and Stiglitz, 1970) of a baseline climate state, as shown in (Baker and Solak, 2011).

The number of technologies, sub-technologies, random technological states of the world, random climate damages states of the world, and the portfolio optimization makes this a high dimensional and large scale computation problem. This makes the problem very computationally expensive to solve in one optimization step within any integrated assessment model (IAM). As such we divide the optimization process into sequential stages, as shown in Figure 9. The specific approach taken is given below.
4.2.2 Overview of Modeling Approach

To model the uncertainty on the R&D process and derive the probability distributions over the scenarios, we rely on prior work by Baker et al. (Baker, Chon et al. 2008, Baker and Keisler 2011, Baker, Chon et al. 2009, Baker, Chon et al. 2008, Baker, Chon et al. 2009, Baker, Chon et al. 2010), who elicit, from top energy experts, the probability of specific levels of R&D funding leading to specific R&D targets for the set of clean energy technologies considered. We note that these R&D targets represent significant costs and performance improvements over the status quo projected development without the stated investment in R&D.

We build on previous work by Zdybel (Zdybel 2013) who modeled all possible energy scenarios that can result from R&D investment in the sub-technology projects within the technologies considered. The impact of the technological scenarios are modeled in a large scale technologically-detailed Integrated Assessment Model, the GCAM. Specifically, Zdybel modeled the impact of each technological scenario on the marginal abatement cost curve (MAC), where the MAC is the cost of reducing greenhouse gas emissions by one additional ton.

In the second step of the sequence (Olaleye and Baker, 2015), we evaluate each of the 3,780 MAC’s corresponding to the different scenarios in a small scale IAM, the Dynamic Integrated Climate-Economic model (DICE). To account for the uncertainty in the impact of climate damages, we reformulate the DICE model (similar to Baker & Solak (2011)) to integrate the impact of stochastic climate damages. The small scale of the DICE IAM allows us to perform large scale scenario analysis, by modeling all combinations of the stochastic damages and the level of technological development of the
scenarios. The DICE model enables us to evaluate the impact of technological advancement in the various energy scenarios in terms of the economic utility to the society.

The third stage of our approach is the R&D investment portfolio optimization. We develop a portfolio allocation model based on a genetic algorithm. The scale and non-differentiability of the problem prevents the use of either exhaustive search or derivative based optimization approaches. We identify the optimal portfolio for a number of R&D budget amounts, under both climate damages and technological uncertainty. We examine the optimal portfolios obtained to determine if they are robust to uncertainty in climate damages and if the portfolios are robust to changes in the constituent technologies.

We also propose a methodology for solving sub-problems of large scale optimization problems. We approach this by decomposing the large portfolio problem and showing that the optimality conditions for the sub-problems are the same as for the large portfolio problem, provided no interactions exists between the constituting energy technologies. We suggest that this methodology is an important step towards dynamic constrained portfolio optimization, given the similar characteristics of sub-compartmentalizing and incrementing portfolio problems.
Figure 9: R&D Portfolio Solution Overview.

This shows the solution approach across the different phases consisting of energy scenarios modeling in GCAM, the climate damages resolution in the DICE model and the technological uncertainty integration in a portfolio optimization model.

The remainder of this paper is organized as follows: In Section 4.3 we detail the problem formulation and solution methodology. In Section 4.4 we provide the results from the study and discuss our findings. Finally in Section 4.5 we summarize the key results, discuss possible future work and conclude the chapter.

4.3 Methodology

The methodology section is divided into four major subsections. In section 4.3.1 we describe the mathematical formulation for the portfolio optimization problem. In section 4.3.2 we discuss the solution methodology used. In section 4.3.3 we discuss the methodology for selecting the budget levels via a greedy heuristic. Section 4.3.4 details the methodology for solving sub-compartment problems of the portfolio problem.

4.3.1 Problem Formulation

In section 4.3.1.1, we detail the mathematical formulation for the portfolio optimization problem; in section 4.3.1.2, we discuss the evaluation of the expected utility of an energy scenario over the stochastic climate damage states.
4.3.1.1 - R&D budget portfolio optimization

For any given R&D budget, our goal is to obtain the optimal allocation of R&D funding to the different energy technologies. That is, we are searching for the optimal portfolio, which maximizes the expected societal utility under uncertainty about climate damages and technological change. The problem formulation is given in Equations (15)-(17) and builds on prior work by Baker and Solak (2014) and Olaleye and Baker (2015).

\[ \max_x \left[ \sum_{\zeta} (p_x(\zeta) \cdot E_\omega(U(\zeta, \omega))) \right] \]  
\[ \text{st.} \quad \sum_{y} \sum_{k} y_{yk} x_{yk} \leq B \]  
\[ \sum_{k} x_{yk} = 1 \quad \forall ij \]

The objective function (Equation (15)) is to maximize the expected societal utility $U$ by choosing the R&D portfolio, $x$, over the different R&D portfolios that can result. $x$ is a matrix, which contains the variables $x_{ijk}$, with each variable equal to 1, if sub-technology $ij$ is funded at level $k$, and 0 otherwise, where $i$ represents the technology category, $j$ represents the sub-technology project and $k$ indicates the level of investment at any of four possible funding levels: no investment, low, medium, or high investment.

The uncertainty in the problem is contained in two random variables, $\omega$, which represents the uncertainty in future climate damage and $\zeta$, which represents the uncertainty in technological development. Each specific portfolio $x$ is associated with a probability distribution, $p_x$, over the possible energy scenarios, $\zeta$. As a reminder, an energy scenario $\zeta$ is defined as a specific set of cost and technical performance specifications for each of the 6 energy technologies considered. The expectation operator $E$ refers to the
expectation over the uncertain climate states, $\omega$. Societal utility $U$ depends on the set of technologies available in the economy and on the climate damages.


Equation (16) is the R&D budget constraint. The equation ensures that the total R&D funding allocated to all the sub-technologies does not exceed the total R&D budget available, $B$. $Y_{ijk}$ is the R&D investment amount for the $k^{th}$ funding level of sub-technology project $ij$.

The constraint in Equation (17) ensures that a sub-technology is invested in at exactly one of the four specified R&D funding levels.

4.3.1.2 Evaluating expected utility of an energy scenario over the stochastic climate damages

In this section, we provide a brief overview of the evaluation of the expected utility $E_{\omega}(U(\zeta, \omega))$ of a specific energy scenario as shown in (15). This is the expected utility given scenario $\zeta$. We evaluate this expected utility in a stochastically modified version of the DICE model\(^5\). The DICE model is a top down economic growth integrated

---

\(^5\) For a complete review of the stochastic reformulation of the DICE model, see (Olaleye & Baker, 2015) and (Baker & Solak, 2014).
assessment model of the climate system and the economy. Developed by Professor William Nordhaus, the model has been widely used in several climate change assessment studies. It has a major advantage in that all the model interactions are expressed in a limited set of equations. The model links the detrimental effect of carbon emissions to economic output through its effect on surface temperature; by balancing capital consumption and savings with the cost of emissions abatement, to determine the optimal tax that maximizes the utility to the society (Nordhaus, 2008; Newbold, 2010). The model aggregates the periodic utility over a 600 year time frame to determine the utility of a given energy scenario. It does this based on a pure rate of time preference discount factor to balance the relative valuation of consumption by present generations compared to consumption by future generations.

While the limited scale of the DICE model makes it amenable to large scale scenario analysis, it has some limitations, including deterministic climate damages and a lack of technological detail. We use a modified version of the DICE model to incorporate stochastic climate states and technological advancement. We present four key equations of the reformulated model: the objective function Equation (18), the abatement cost function, Equation (19) and non-anticipativity constraints (20) and (21). R&D-induced technological advancement is modeled to kick in at 2055, and the uncertainty about the severity of climate damage is also assumed to be resolved by 2055. Therefore, while decisions are made under uncertainty about climate damages during the 2005 – 2055 stage, the resulting climate damage levels are explicitly known by 2055.

\[
\max_{\mu^{i}, y^{i}} \sum_{o \in \Omega} p^{o} \sum_{t} R^{o}_{t} u^{o}_{t}
\]  

(18)
Equation (18) maximizes the discounted sum of per capita utility of consumption, \( u_t \). As in Equation (15) \( \omega \) denotes a possible future climate state of the world and \( \Omega \) is the set of all possible climate states. \( p^{\omega} \) is the probability of climate state \( \omega \). \( u_t \) is the per-period utility and \( R_t \) indicates the discount factor for period \( t \). The two key decision variables in this model are abatement – a reduction in emissions below a business as usual level – and the standard investment in capital. Equation (19) is the cost of abatement, \( \mu \). It has been modified from original DICE model. \( \mu_t \) is the fraction of emissions abated in period \( t \), \( \pi_t \) is a regional participation factor (related to how many countries cooperate in abatement), \( \theta_1 \) and \( \theta_2 \) are the abatement cost function parameters and \( \zeta \) is an energy scenario. Each technology scenario \( \zeta \) is associated with a set of parameters \( \alpha_0^{\zeta} \) through \( \alpha_3^{\zeta} \), which impact the cost of carbon abatement function as shown in Equation (19). The derivation of Equation (19) and the calibration of the \( \alpha_0^{\zeta} - \alpha_3^{\zeta} \) parameters are discussed explicitly in Olaleye & Baker (2015). Given that the formulation is a sequential stochastic programming problem, equations (20) and (21) enforce non-anticipativity of the first stage decision variables in the formulation; this ensures that decisions taken in the first stage (pre 2055) are the same for all the
realizations of the climate damages state. Baker & Solak (2013) show that these two constraints are sufficient to enforce non-anticipativity in the DICE model.

4.3.2 Solution Approach: Genetic Algorithm

Several optimization techniques exist for solving problems with large search spaces; some of these include but are not limited to gradient descent methods, pattern search, swarm intelligence, evolutionary algorithms, branch and bound, cutting plane, simulated annealing, co-ordinate search and Gaussian adaptation. The peculiar nature of our R&D portfolio optimization problem – non-convexity, non-differentiability of the objective function and constraints, integer constraints and the large search space – limits the use of most of these optimization approaches. Swarm intelligence and evolutionary meta-heuristic algorithms have been shown to be well suited to optimization problems of this nature (Kennedy, Kennedy, Eberhart, & Shi, 2001). Based on this, the optimization approach we use is the genetic algorithm (GA) hybridized with the simulated annealing algorithm. A GA (Goldberg, 1989) is a meta-heuristic search approach that mimics nature’s adaptation strategy by combining a survival of the fittest strategy with some degree of luck (randomness in survival). Gas are intuitive and the flexible as the properties of the algorithm, e.g. cross-over probability and hybridization, can be adjusted over the course of the solution to aid convergence depending on the solution stage. This has made it a widely used optimization approach for problems with very large solution spaces (Goldberg, 2002; Hassan et al., 2005). In subsequent sections below, we discuss our approach to implementing the optimization technique to our R&D portfolio problem. In section 4.3.2.1, we discuss the specifications of the algorithm, including the seeding criteria, offspring selection, mutation and the solution convergence criteria of the
algorithm. Complexity reduction techniques, hybridization and the implementation of the algorithm are discussed in section 4.3.2.2.

### 4.3.2.1 Fitness Function, Seeding, offspring selection and convergence criteria

The GA algorithm operates by iteratively modifying and improving current feasible solutions, similar to the natural selection hereditary process. Each feasible candidate solution, a funding portfolio in our portfolio optimization problem, can be represented through a chromosome representation. A chromosome corresponds to a portfolio, a specific set of funding levels for all the sub-technologies, in our particular problem. E.g. \( \bar{Y} = [4,1,2,1,2,3,4,3,2,1,3,4,2,3,3,2] \) where 1 is the high funding and 4 is no funding. The bits of the chromosomes represent the recommended funding level for each of the sub-technologies.

Similar to a hereditary process, the subsequent generations’ are offspring’s of the prior generations, \textit{parents}. These offspring’s are created by modifying the bits of the chromosomes of the parents, through a combination of mutation, cross-breeding and inheritance techniques. Mutation indicates that the new offspring’s chromosome bit is a slight modification to a single parent’s; cross-breeding ensures that the new offspring’s chromosome bit is a blend of the parents’; and inheritance involves the direct passing over of an exact chromosome gene to the offspring from a parent. On each iteration of the algorithm, feasible candidate solutions with a higher rank, in terms of the objective value, have a higher probability of being selected. To evaluate the objective value of these candidate solutions, a fitness function is used.
**Fitness Function:** The fitness function of the GA evaluates the objective function for a specific candidate solution. It evaluates the expected utility of a specific R&D portfolio, over all the energy scenarios that can result. This is because the uncertainty in the R&D process can lead any funding portfolio to stochastically lead to different combinations of R&D targets leading to different energy scenarios.

**Seeding:** The seed is the feasible candidate solution, or chromosome, the algorithm starts from. The initial implementation of the algorithm starts by randomly generating a feasible solution (any portfolio). Other implementations of the algorithm were also carried out by starting with the best current solution from a previous implementation of the GA algorithm or with the solution from a best fist greedy algorithm discussed later in section 4.3.3.1. These implementations were made to explore the robustness of the GA solution and create the R&D budget levels respectively.

**Evolution, Selection:** Chromosomes of the starting population are randomly mutated, crossbred or passed-over to generate the next generation of candidate solutions. Candidate solutions are ranked in order of their fitness function value. Higher probability of selection is assigned to higher ranked candidates. Of these current populations, the offspring with the highest rank, elite candidate, in terms of the fitness function value, have a higher probability of being selected with a condition to enforce diversity of the offspring. The elites are initially constrained to be distinct from each other; this is relaxed as the algorithm progresses, similar to a simulated annealing process. The process is repeated until the maximum iteration is reached or the change in the objective function falls below a defined threshold. The candidate solution with the best fitness function value at the end is selected as the best solution.
4.3.2.2 Implementation

We run the algorithm with different specifications of the starting solution, population size, cross-over probability, mutation probability, maximum iteration and minimum iteration. The algorithm is run a minimum of 10 times for each budget level considered, as is best practice (Mitchell, 1996; Eiben et al., 1999). We show the different specification parameters of the algorithm in Table 6. The different possible initialization of the GA are given in rows three to five of the first column of the table. A random GA represents a form of the GA with the initial GA starting solution randomly selected, the best current solution GA is a form of the GA where the initial GA starting solution is the solution of a previously conducted random GA and the best fit greedy GA represents when the initial GA starting solution is the best from a previously conducted greedy best fit algorithm, discussed in section 4.3.3. We implement the GA primarily using the random starting solution with different combinations of the other properties given in columns two to seven. In the event that there is an obvious need for improvement in the GA solution, we implement using the best current solution GA. In addition, we also compare our GA implementation with the built-in genetic algorithm toolbox in Matlab (The MathWorks 2012) using the different combinations of the rank scaling, selection function, mutation and crossover parameters. We find no difference between these implementations.

Table 6: Parameters of the Genetic Algorithm

<table>
<thead>
<tr>
<th>Initial Population</th>
<th>Population Size</th>
<th>Selection</th>
<th>Mutation</th>
<th>Crossover</th>
<th>Stall Iteration</th>
<th>Maximum Iteration</th>
</tr>
</thead>
</table>

72
<table>
<thead>
<tr>
<th>Random</th>
<th>Best Current Solution</th>
<th>Stochastic</th>
<th>6%</th>
<th>50%</th>
<th>50</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td></td>
<td>100</td>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>200</td>
<td>15%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Genetic Algorithm Implementation

<table>
<thead>
<tr>
<th>No of Systems</th>
<th>Solution Time (hours)</th>
<th>Minimum Runs/Budget Level</th>
<th>Budget Levels</th>
<th>Risk Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel</td>
<td>30</td>
<td>3 - 14</td>
<td>10</td>
<td>31</td>
</tr>
</tbody>
</table>

Scale of portfolio problem. As most of the 17 sub-technologies are elicited at 4 R&D funding levels, some of which can succeed at different R&D targets, the problem space is large. There are 2,783,138,807,808 possible optimal funding portfolios, each of which maps the possible set of 2,239,488 combinations of R&D targets to the 3,780 energy scenarios. The evaluation of the fitness function for a single candidate solution is expensive as it involves mapping the different 2,239,488 R&D outcomes to each of the 3,780 energy scenarios to assess the probability distribution resulting from the specific candidate funding portfolio. To exhaustively compute an optimal solution using exhaustive search for a single budget level would require over a year even with a 100 computation nodes.

Complexity Reduction. Naïve implementation of the fitness function of the GA algorithm as discussed in the formulation section would take on average over 400 seconds for the evaluation of the fitness evaluation of a single candidate portfolio, and a computation time of over 23 days for the evaluation of the GA algorithm, at a modest 50 maximum iterations with a population size of 100. We reduced the evaluation time of the fitness function of the algorithm by pre-processing the societal utilities of the energy scenarios and then also pre-assigning the societal utility for the energy scenarios to all the possible sets of R&D combinations of R&D targets that can result. As such we pre-map
all the 2,239,488 combinations of R&D targets to each of the energy scenarios; this ensures that the GA algorithm does not have to dynamically map these scenario utilities to the R&D outcomes each time the fitness function for a portfolio is evaluated. Given the pre-processing and pre-assignment of the utilities, the fitness function now only computes an expected value using the probability distribution over the scenarios, where the distribution is a function of the funding portfolio currently evaluated. Along with several other improvement implementations, we reduced the average computation time to around 9 seconds per fitness evaluation, leading to an average GA evaluation time of 9 hours. We run the algorithm in a cluster of 30 quad core systems with the details of the implementation shown in Table 6.

4.3.3 Choosing Budget Levels

The size of the total R&D budget available is a major determining factor in choosing the optimal allocations to the different energy technologies. To avoid randomly picking budget levels to optimize over, we utilize a greedy heuristic to determine the appropriate R&D budget funding levels. The algorithm is discussed below in Section 4.3.3.1.

4.3.3.1 Best fit Greedy Heuristic Algorithm

We implement the best fit greedy algorithm to determine the appropriate R&D budget funding levels. A greedy algorithm is an algorithm that chooses the optimal value at each subset of the problem without considering the other subsets of the problem when solving a problem composed of several subsets (Cormen et al., 2001). To apply the greedy algorithm, we first classify all the R&D targets that can result at different levels of
R&D funding, we denote each unique combination of an R&D target and a funding level as an R&D-target-to-funding-level combination. We then calculate an R&D effectiveness ratio for all of these R&D targets that can result given different levels of funding. This ratio is obtained by dividing the expected utility of an R&D target when invested in at a particular funding level by the probability of success, as shown in Table 7. Finally, we order all the R&D-target-to-funding-level combinations in terms of their R&D effectiveness ratio and combine them together in descending order of their R&D effectiveness ratio.

The Greedy algorithm works by ranking the R&D targets of the R&D-target-to-funding-level combinations, in decreasing order of their R&D effectiveness, starting from the combination with the highest R&D effectiveness ratio. The funding levels corresponding to these R&D-target-to-funding-level combinations are then cumulatively added to generate successive R&D budget levels. We approximate the R&D budget levels obtained to the nearest hundred million dollars for tractability. This resulted in 31 unique budget levels as shown in Figure F.1.

For the approach to work, we ensure some constraints hold. The first constraint is that we ensure that only incremental R&D funding is added to the budget, for an R&D target corresponding to the same sub-technology project that has had another R&D target previously considered. The second constraint is that if the R&D-target-to-funding-level combination to be added to the portfolio has a lesser funding level than a previously considered R&D-target-to-funding-level combination, no R&D funding is added. We also ensure that, if the R&D-target-to-funding-level combination to be added to the portfolio is for a different success level of the R&D target of a previously added R&D-
target-to-funding-level combination, no additionally R&D funding is added to the greedy portfolio. A major disadvantage of this method is that it does not factor in relationships between the energy technology groups.

Table 7: R&D effectiveness of the sub-technologies

Assessed at the medium R&D funding level, the sub-technology high success endpoint and the No Damage Uncertainty Case (Olaleye and Baker, 2015).

<table>
<thead>
<tr>
<th>Technology</th>
<th>Sub Technology</th>
<th>Weighted Funding [Funding/Success Probability] ($million)</th>
<th>Utility</th>
<th>R&amp;D Effectiveness [Utility/Weighted Funding]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon Capture &amp; Storage</td>
<td>Post-Combustion</td>
<td>262</td>
<td>53</td>
<td>0.2016</td>
</tr>
<tr>
<td>Solar</td>
<td>In-Organic</td>
<td>266</td>
<td>13</td>
<td>0.0505</td>
</tr>
<tr>
<td>Carbon Capture &amp; Storage</td>
<td>Chemical Looping</td>
<td>280</td>
<td>76</td>
<td>0.2732</td>
</tr>
<tr>
<td>Bio-Electricity</td>
<td>Steam</td>
<td>414</td>
<td>5</td>
<td>0.0129</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Hydrolysis</td>
<td>427</td>
<td>2</td>
<td>0.0037</td>
</tr>
<tr>
<td>Nuclear</td>
<td>Light Water Reactor</td>
<td>769</td>
<td>122</td>
<td>0.1590</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Gasification</td>
<td>1,242</td>
<td>7</td>
<td>0.0055</td>
</tr>
<tr>
<td>Bio-Electricity</td>
<td>Gasification</td>
<td>1,268</td>
<td>9</td>
<td>0.0073</td>
</tr>
<tr>
<td>Carbon Capture &amp; Storage</td>
<td>Pre-Combustion</td>
<td>1,364</td>
<td>65</td>
<td>0.0478</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Sel. Thermal Proc. 1</td>
<td>1,514</td>
<td>2</td>
<td>0.0015</td>
</tr>
<tr>
<td>Batteries for Elect. Trans.</td>
<td>Lithium-Ion</td>
<td>1,691</td>
<td>4</td>
<td>0.0023</td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Sel. Thermal Proc. 2</td>
<td>2,364</td>
<td>1</td>
<td>0.0003</td>
</tr>
<tr>
<td>Batteries for Elect. Trans.</td>
<td>Lithium-Metal</td>
<td>3,472</td>
<td>5</td>
<td>0.0014</td>
</tr>
<tr>
<td>Nuclear</td>
<td>High Temp Reactor</td>
<td>11,209</td>
<td>122</td>
<td>0.0109</td>
</tr>
<tr>
<td>Solar</td>
<td>Organic</td>
<td>24,692</td>
<td>13</td>
<td>0.0005</td>
</tr>
<tr>
<td>Nuclear</td>
<td>Fast Reactors</td>
<td>926,600</td>
<td>122</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
4.3.4 Sub-Compartmentalizing the Portfolio Allocation Problem

In this section, we discuss our methodology to decomposing constrained portfolio allocation problems with a view to solving sub-problems of the portfolio optimization problems. In section 4.3.4.1 we provide the formulation for sub-compartmentalizing our R&D portfolio optimization problem, in Section 4.3.4.2 we propose the decomposition method, in Section 4.3.4.3 we apply the decomposition methodology to sub-problems of large scale problems, and finally in Section 4.3.4.4 we summarize the decomposition methodology.

4.3.4.1 Sub-Compartmentalizing the R&D Portfolio Allocation Problem

In this section, we discuss our approach for solving sub-problems of our R&D portfolio problem. Solving the sub-problem of constrained optimization problems is interesting for several reasons. First, it enables the possibility of carrying out sensitivity analysis to examine the robustness of the optimal portfolio. As an example, we can examine the effect of removing CCS from the portfolio, based on the relationship of CCS with others technologies considered. Additionally, we feel that ultimately such decomposition techniques can be extended to solving large scale intractable problems.

The formulation of the sub-problem constrained optimization problem is a slight modification to Equations (15) – (17) and is given below in Equations (22) – (24). \( \hat{x}_{ij,k} \) indicates which level of funding is allocated to the excluded sub-technology project, \( \hat{\zeta} \), in the optimal portfolio. We note that the optimal portfolio \( x \) for the sub-compartment portfolio problem does not include sub-technology project \( \hat{\zeta} \).

\[
\max_{x} \left[ \sum_{\zeta} \left( p_{x}(\zeta) \ast E_{\omega} (U(\zeta, \omega)) \right) \right] \tag{22}
\]
\[ \sum_{j_y j \neq j} \sum_{k} Y_{jyk} x_{jyk} \leq B - \left[ \sum_{j} \sum_{k} Y_{jyk} x_{jyk} \right] \] (23)

\[ \sum_{k} x_{jyk} = 1 \quad \forall ij : ij \neq \tilde{ij} \] (24)

Equation (23), is modified so that the new budget is the previous budget \( B \) less the previously allocated optimal funding to the omitted sub-technology \( \tilde{ij} \).

### 4.3.4.2 Decomposition Methodologies for Portfolio Optimization problems

Any constrained portfolio optimization problem can be expressed as the maximization of an objective \( u(x) \) subject to a budget constraint \( B \), given in equation (25) below.

\[
\begin{align*}
\max_{x} u(x) \quad & s.t. \quad \sum_{j} \sum_{k} Y_{jyk} x_{jyk} \leq B \\
\end{align*}
\] (25)

As an example, \( u(x) \) can be defined as shown in Equation 15 as

\[ u(x) = \sum_{\zeta} \left( p_{\zeta} (\zeta) \ast E_{\omega} \left( U(\zeta, \omega) \right) \right) \] and \( x \) is defined as in Section 4.3.1.1. \( \hat{x} \) is the portfolio that maximizes the portfolio optimization problem, defined as

\[ \hat{x} = \arg \max_{x} u(x) \] (26)

We write the optimized function as \( u(x)|_{x = \hat{x}} \).

Here we introduce a notation for denoting sub-sets of a portfolio. A portfolio sub-set is defined as a portfolio where some technologies are \textit{a priori} excluded from consideration from the set of feasible technologies. To properly denote this, we introduce a new variable, \( \chi_i \), which is a variable denoting the R&D investment amount in a specific technology \( i \). For example, \( \chi_i \) represents the R&D investment amount in
technology 1 and is defined as \( \chi_i = \sum_j \sum_k \gamma_{1jk} x_{ijk} \), with \( x_{ijk} \) as previously discussed a binary variable denoting the investment decision at a specified funding level \( k \) of a specific sub-technology project \( ij \).

To denote a sub-set portfolio, we omit the \( \chi_i \) corresponding to the excluded technology. For example, the portfolio vector \( \left( \chi_i, \chi_j, \chi_k, \chi_l \right) \) indicates that only four technologies \( i, j, k, l \), are considered in the portfolio optimization problem and all other technologies are a priori excluded \( \left( \chi_i, \chi_j, \chi_k, \chi_l, 0, \ldots, 0 \right) \), with their R&D funding amount set to 0.

**Portfolio decomposition.** Note that the original portfolio optimization problem \( \max_x u(x) \) can be decomposed as follows:

\[
\max_x u(x) = \max_x \left[ \sum_i f_i(\chi_i) + \sum_{ij} f_{ij}(\chi_i, \chi_j) + \sum_{ijk} f_{ijk}(\chi_i, \chi_j, \chi_k) + \ldots \right]
\]

(27)

where the first order term \( f_i(\chi_i) \) is defined as: \( u_i(\chi_i, 0, 0, \ldots, 0, 0) \). That is, it is the payoff from having a portfolio with R&D investment \( \chi_i \) in a specific technology \( i \), and an investment of 0 in all other technologies.

The second term \( f_{ij}(\chi_i, \chi_j) \) is defined as
\[ f_{i,j}(\chi_i, \chi_j) = u_{i,j}(\chi_i, \chi_j) - f_i(\chi_i) - f_j(\chi_j) \]  

(28)

Where \( u_{i,j}(\chi_i, \chi_j) \) is the payoff from having investment \( \chi_i \) and \( \chi_j \) in technologies \( i \) and \( j \) respectively, and an investment of 0 in all other technologies. Equation 28 reflects the additional payoff from the interaction of the two technologies above the value of the individual technologies alone.

Higher order terms can be expressed similarly, as shown by the form

\[
\begin{align*}
f_{i,j,k,\ldots,N}(\chi_i, \chi_j, \chi_k, \ldots, \chi_N) &= u_{i,j,k,\ldots,N}(\chi_i, \chi_j, \chi_k, \ldots, \chi_N) \\
&- \left[ u_{i,j,k,\ldots,N-1}(\chi_i, \chi_j, \chi_k, \ldots, \chi_{N-1}) - u_{i,j,k,\ldots,N-2}(\chi_i, \chi_j, \chi_k, \ldots, \chi_{N-2}) - \ldots - \sum_{i}^{N} f_i(\chi_i) \right] \\
&- \left[ - \ldots - \sum_{i \neq j}^{N} u_{i,j}(\chi_i, \chi_j) - f_i(\chi_i) - f_j(\chi_j) - \sum_{i}^{N} f_i(\chi_i) \right]
\end{align*}
\]  

(29)

4.3.4.3 Sub-problem optimization for sub-compartmented problems

Case I: No second or higher order interactions between all constituent technologies

We define a no interaction case between constituents technologies as when all higher order factors between the technologies are zero. In this case the objective function is as shown in Equation 30.

\[ \max_{x} u(x) = \max_{x} \left[ \sum_{i}^{N} f_i(\chi_i) \right] \]  

(30)

Case II: No interaction between constituent technologies in an unconstrained optimization problem.
The solution here is trivial as the optimization problem reduces to Equation 31 shown below and the subproblem $\max_{x: u(x)} u(x)$ results in the same portfolio as

$$\sum_i \max_x u_i(x_i).$$

$$\max_x \left[ \sum_i f_i(x_i) \right] = \max_{x_1} \left[ f_1(x_1) \right] + \max_{x_2} \left[ f_2(x_2) \right] + \ldots + \max_{x_n} \left[ f_n(x_n) \right]$$  \hspace{1cm} (31)

Case Iii: No interaction between constituent technologies in a constrained optimization problem.

For simplicity, consider a 3 technology constrained problem as shown in Equation 32.

$$\max_x f_1(x_1) + f_2(x_2) + f_3(x_3) \quad s.t. \quad x_1 + x_2 + x_3 \leq B$$  \hspace{1cm} (32)

Assuming this problem is solved to optimality and we find no pairwise or higher order interaction between the 3 technologies. The sub-problem formed by removing the optimal funding for the third technology, $x_3$, is shown in Equation 33.

$$\max_x f_1(x_1) + f_2(x_2) \quad s.t. \quad x_1 + x_2 \leq B - x_3^*$$  \hspace{1cm} (33)

We can set this up as a constrained lagrangian problem as shown in Equation 34 and solve for the first order conditions of the sub-problem as shown in Equation 35.

$$f_1(x_1) + f_2(x_2) = \lambda \left( x_1 + x_2 - B + x_3^* \right)$$  \hspace{1cm} (34)
\[
\frac{\partial u(x)}{\partial \chi_1} : f_1' - \lambda = 0
\]
\[
\frac{\partial u(x)}{\partial \chi_2} : f_2' - \lambda = 0
\]
\[
\frac{\partial u(x)}{\partial \lambda} : \chi_1 + \chi_2 + \chi_3 - B \leq 0
\]  

(35)

We observe that this is the same first order conditions as for the original 3 -technology problem as shown in Equation 36.

\[
\frac{\partial u(x)}{\partial \chi_1} : f_1' - \lambda = 0
\]
\[
\frac{\partial u(x)}{\partial \chi_2} : f_2' - \lambda = 0
\]
\[
\frac{\partial u(x)}{\partial \chi_3} : f_3' - \lambda = 0
\]
\[
\frac{\partial u(x)}{\partial \lambda} : \chi_1 + \chi_2 + \chi_3 - B \leq 0
\]

(36)

This shows that both problems are equivalent.

4.3.4.4 Empirical Expectation based on theoretically finding when the sub-problem optimization is applied to R&D portfolio problem

In Olaleye and Baker, (2015), we estimated the interaction coefficients between the constituent technologies in our R&D portfolio problem based on a regression methodology. Figure 5 shows the results. We hypothesize that the optimal solution for a decomposed problem will be nearly identical to the optimal solution for the complete problem when the interaction coefficients are small and/or insignificant, such as for Batteries for Electric Transportation, where no interactions exists with any other energy technology.
**Methodology description summary.** We note that our procedure above can be described as a case of a backward recursive algorithm. One solves a large scale problem, and then examines if any technology exists that has no interactions with other constituent technologies. If this exists, then the solution to the sub-problem is available from the original problem without a need for any optimization. This can iteratively go on to smaller sub-problems.

### 4.4 Results

In this section we discuss the results from our study. In section 4.4.1, we outline, in detail, the insights about the composition of the optimal portfolio when there is no uncertainty about climate damages (the No Damage Risk case) while section 4.4.2 espouses the additional insights that arise in the composition of the optimal portfolio when there is uncertainty in the impact of climate damages. Section 4.4.3, highlights the results from the solving of sub-compartments of large portfolio problems.

#### 4.4.1 Optimal Portfolio in the No Damage Risk case

As discussed in section 4.3.3.1, the funding budget levels are chosen by cumulatively summing the budgets of the sub-technologies, where the sub-technologies have been initially ranked by their R&D utility to funding cost ratio.

The optimal portfolios in the No Damage Risk case are shown in the appendix in Table F.1. The table shows the recommended allotment of funding for each of the sub-technologies and the corresponding utility for each funding portfolio.
Figure 10 and Figure 11 show the allocations of R&D funds to these technologies with the corresponding societal utility gain for each budget levels, with Figure 10 showing the R&D levels for budgets up to the $6.1 billion levels for a closer view and Figure 11 the allocation for all the budget levels.

**Investment Priority for CCS and some specific sub-technologies in Solar and Nuclear:** We observe from Figure 10 that the optimal portfolio prioritizes investment in CCS sub-technologies, Nuclear Light Water Reactor sub-technology and the Solar Inorganic sub-technology at any given budget level.

**Diminishing Returns to R&D:** From Figure 10 and Figure 11 we find that the optimal portfolios exhibit diminishing returns to scale especially after the $6 billion R&D budget when all CCS and Batteries for Electric Transportation technologies are fully invested in. This reinforces the need for portfolio optimization, as 20% of the total possible funding amount can result in 90% of the utility gain.

**Nuclear Investments are robust and favorable but require very large R&D funding:** We find that Nuclear sub-technologies perform relatively well, but we note that they require significantly larger R&D funding amounts (Olaleye & Baker, 2015). A look at Figure 10 and Figure 11 also show that Nuclear investments are robust as the allocation to Nuclear never decreases with an increase in the R&D budget level.
Figure 10: Optimal R&D Portfolio to Utility chart for the $300 million to $6.1 billion R&D budget levels when there is no uncertainty in climate damages.

Figure 11: Optimal R&D Portfolio to Utility chart across all the R&D budget levels when there is no uncertainty in climate damages.
4.4.2 Effect of Climate Damage Risk on the Optimal Portfolio

In this section, we examine the effect of risk in climate damages on the optimal R&D portfolio where the damage risk cases are defined as mean preserving spreads of the No Damage Risk case as shown in the appendix in Table D.1. Figure 12 shows the additional societal utility gain of the R&D portfolios, due to technological advancement, over the three risk cases. The societal utilities shown in the figure represent the additional gain in utility of each of the R&D portfolio for each of the three risk cases, with respect to the No technological advancement scenario in each of the risk cases. Figure 13 shows the change in the optimal portfolio of the Medium and High Damage Risk cases with respect to the No Damage Risk optimal portfolio. The optimal portfolio for the Medium and High Damage Risk case are provided in the appendix.

Impact of R&D is greatest in the Medium risk case: We find that the effect of technological advancement is not monotonic in damage risk. As observed in previous papers (Baker & Solak, 2011), (Olaleye & Baker, 2015), from Figure 12, technological advancement has more value in the Medium risk case, followed by the No Damage Risk case and the High risk case respectively. This is because in the Medium risk case energy technologies choose an optimal level of carbon emissions to abate in addition to reducing the cost of abatement ‘cost effect.’ In the High damage risk case, however, in the event that climate damages do occur (5% probability), the technologies are already constrained to fully abate carbon emissions and can only reduce the cost of carbon emissions abatement.

Increase in climate damage risk favors investment in CCS and Bio-Fuels and dis-incentivizes Solar and Batteries for Electric Transportation investment: From Figure 13 we find that higher climate damages risk favors investment in CCS until the
maximum CCS investment is reached at the expense of investment in Solar and Electricity from Batteries. We also find that, after the maximum CCS investment is reached, there is an incentive for investment in the Bio-Fuels technology in lieu of Batteries for Electric Transportation. The desirability of CCS and Bio-Fuels is due to their marginal abatement cost MAC curves (shown in Figure 14) with the two technologies having lower full, 100%, abatement costs than their competing counterparts. CCS has lower full MAC than all technologies, with its abatement cost generally lower than all technologies apart from Nuclear, and the scale of R&D investment needed is relatively very low. Bio-Fuels MAC is initially higher than other competing technologies such as Batteries for Electricity Transportation and Solar energy but becomes cheaper comparatively (Figure 15) as the level of abatement increases. For solar technology, we see that, though it performs relatively well compared to Bio-Electricity at lower abatement levels, this relationship reverses as we approach full carbon abatement. This reinforces the desirability of having development in CCS and Bio-Fuels as they can serve as hedges in the event that higher than anticipated climate damages occur (Olaleye & Baker, 2015).

**Investment in Nuclear technology is invariant to damage risk:** We note that R&D investments in Nuclear technology are robust to climate risk. This is because Nuclear technologies are capital intensive but have a strong impact on the abatement cost curve (Figure 14). Therefore Nuclear technology is a relatively constant inclusion in the optimal portfolio provided there are enough funds for its development; it does not however provide any added benefits in hedging against climate risk. The investments in Nuclear are monotonically increasing.
4.4.2.1 Robustness of optimal portfolio to climate damage risk.

We conducted an analysis to evaluate the frequency of having the same portfolio optimal for two or more risk cases. We find that the optimal portfolio varies significantly with climate risk. Of the 31 budget levels considered, the optimal R&D portfolio was the same for all the three damage risk cases in only 9 of the 31 budget levels. We also note that higher damages risk cases tend to have similar optimal portfolios as Medium and High risk cases share 9 optimal portfolios while the same portfolio is not optimal for the No and High risk cases. 16 of the 31 optimal portfolios in the No Damage Risk case are not the same with those from any of the other two risk cases, while only 7 and 13 such
portfolios exist in the Medium and High Damage Risk cases, respectively. In summary, we find that climate damage risk plays a significant part in the optimal portfolio.
Figure 13: Change in Optimal Portfolio with Damage Risk with respect to the No Damage Risk (Med Risk = Medium Damage Risk Portfolio – No Damage Risk Portfolio and High Risk = High Damage Risk Portfolio – No Damage Risk portfolio)
Figure 14: Abatement Cost curves with only advancement in one sub-technology.

Figure 15: Marginal abatement cost curves 60% - 100%.
4.4.3 Sub-compartmentalizing the Portfolio Problem

In this section, we discuss results from solving sub-problem decompositions of our R&D portfolio problem. As discussed in section 4.3.3, we show that the decomposition methodology will result in exact solutions only when there are no interactions between the component energy technologies. The regression shown in Table 3 from Olaleye & Baker (2015) shows that such interactions between technologies exist between most technologies. It also provides a means of quantifying the interaction characteristics between the technologies. Based on this, we consider five sub problems where we respectively remove the optimal allocation for (1) Batteries for Electric Transportation (2) Carbon Capture and Storage, (3) Nuclear, (4) Bio-Electricity and (5) the Carbon Capture and Storage and Bio-Electricity pair. We select this particular set of technologies based on their interaction characteristics with the other technologies. The first case, Batteries for Electric Transportation, has no observed interactions with any of the other technologies, hence sub-compartmentalization should result in the same portfolio as the six technology problem. The other cases, have some significant interactions with other technologies; we include this to show the change in the portfolio. Carbon Capture and Storage is a strong complement for Bio-Electricity and it is substituted for Nuclear and Solar for the No Damage Risk case. Nuclear is a strong substitute for Bio-Electricity, Solar and Carbon Capture and Storage. Bio-Electricity plays a unique role as it is a strong complement for Carbon Capture and Storage while it is also a strong substitute for Nuclear. We also consider the Carbon Capture and Storage – Bio-Electricity pair given that it is the only significant complement. To verify our analysis from section 4.3.4, we solve the above highlighted sub-problems and examine the results with those from the regression result in Table 3. We discuss the Bio-Electricity
case and the CCS case in the sections below and show all other in the Appendix (APPENDIX H).

4.4.3.1 Sub-compartmentalizing based on technologies with no interactions effects with other technologies

In this section, we sub-compartmentalize the R&D portfolio problem based on technologies which have no or insignificant interactions with other technologies. From our decomposition analysis in section 4.3.4, we expect the optimal portfolio to be the same as there is independence between the constituents of the portfolio.

Optimal Portfolio in the 5 technology problem when the optimal allocation to Batteries for Electric transportation is excluded: Similar to the above figures, Figure H.2 shows the difference in the sub-problem optimal portfolio when Bio-Electricity energy technology is excluded. However, we note that the optimal portfolio for the five technology portfolio problem is always exactly the same with that of the larger six technology portfolio problem, Table F.1, are the same for budget levels shown. This also supports findings from section 4.3.4, which show that, when no second or higher order interactions exist between technologies, the optimal portfolio will be the same.

4.4.3.2 Sub-compartmentalizing based on technologies which have interactions with other technologies

In this section, we sub-compartmentalize the R&D portfolio problem based on technologies which have interactions with other technologies. From our decomposition analysis in section 4.3, we expect the optimal portfolio to change in accordance with the nature of the interactions between technologies, as assessed in Table 3.
Change in the Optimal Portfolio in the 5 technology problem with respect to CCS: Figure 16 shows the additional benefit of having CCS in the portfolio. The figure displays the difference between the six technology portfolio and the five technology portfolio. A positive change in the allocation to a technology indicates that the investment in that technology has decreased in the 5-tech port relative to the 6-tech. This indicates that there is a synergistic effect exists between CCS and the technology exhibiting the change; a negative change indicates the converse while no change indicates independence between CCS and the specific technology.

We find that the presence of CCS in the portfolio causes the optimal portfolio to vary as investment in Bio-Electricity technology increases at the expense of Batteries for Electric transportation, and less so for Solar and Bio-Fuels energy technologies. We note that this difference in the portfolio is very significant, ranging from $200 million to almost a $1 billion, emphasizing the need to consider the effect of synergies between technologies when solving the portfolio problem. We note that these changes support the findings from Table 3 that shows that Carbon Capture and Storage and Bio-Electricity are very significant energy complements.
Figure 16: Change in optimal portfolio with respect to CCS.

*It shows the difference in the optimal portfolio between the six technologies problem and the optimal portfolio in the five technologies problem.*

### 4.5 Summary/Conclusions

In this chapter, we discussed the optimal allocation of R&D funds to combat climate change in a portfolio of 6 low carbon energy technologies. We solved the problem using a genetic algorithm, addressing the uncertain climate damages and technological feasibility problem sequentially in two stages. The climate damage uncertainty is resolved in a stochastically reformulated version of the DICE model. We then unraveled the technological uncertainty from R&D funding and optimize over all the climate damage risk cases using a genetic algorithm implementation.

We showed that the optimal allocation of the R&D funds exhibits considerable diminishing returns to scale, especially after the $6-7 billion budget level. This shows a
clear need to address the optimal portfolio problem from an optimization viewpoint. The optimal funding portfolio prioritizes reaching maximum R&D investment early in certain technologies. The most significant technology in terms of its impact is the CCS technology. We also find that Light Water Reactor Nuclear sub-technology and the In-Organic sub-technology are always funded in the portfolio.

We examined the effect of climate damage risk on the level technological advancement. We find that the optimal R&D portfolio varies significantly with climate damage risk, leading to more R&D investment in CCS and Bio-Fuels technologies, and less investment in the Solar, Batteries for Electric Transportation and Bio-Electricity technologies as damage risk increases. We observe that investments in Nuclear technology, however are very robust to climate risk; this is because Nuclear technologies cause significant reduction to the reference MAC, but require very high R&D development costs.

A major problem with energy R&D portfolio optimization is the constantly varying flux of the exciting prospective energy technologies. It is therefore plausible that policymakers might want to remove from consideration a technology (or set of technologies) in a portfolio that is already optimized. For example, suppose the Department of Energy has an already balanced energy investment portfolio for the period 2015-2020, with positive budgets for Nuclear R&D. Yet, after an unexpected major accident like that at Fukushima, it may no longer want to consider development of Nuclear. We show that, in cases like these, there is no need to re-optimize the sub portfolio, provided the excluded energy technology has no interaction with the other technologies and sets of technology in the portfolio. We anticipate that a version of this
methodology may be also extended to constrained portfolio optimization problems where there exists interactions between the excluded technology and other technologies in the portfolio. We conduct an empirical study based on our 6 technology R&D portfolio problem and show that the optimal sub portfolio is the same if there does not exist any interactions with the reference technology removed. We do this by solving different sub-problems of the actual portfolio optimization. We find that when Batteries for Electric Transportation is removed, there is no change in the portfolio, since Batteries for Electric Transportation does not interact with any of the other technologies (or set of technologies). We also find that the exclusion of technologies that have interactions with other technologies (or set of technologies) lead to changes in the sub problem portfolio.
CHAPTER 5

PART 3: GLOBAL UNCERTAINTY ANALYSIS OF THE GLOBAL CHANGE ASSESSMENT MODEL: QUANTIFICATION, MODELING AND ASSESSMENT

5.1 Introduction

This chapter of the thesis is focused on the modeling and characterization of uncertainties in large scale integrated assessment models (IAMs). We present a methodology for conducting global sensitivity analysis (GSA) in a large scale climate IAM, the Global Change Assessment (GCAM) model. GSA describe a broad array of methods for allocating the uncertainty in model outputs to the different constituent model inputs. They differ from traditional sensitivity analysis, in that they consider the entire input parameter space at once, accounting for input parameter interactions and structural model effects, rather than considering one factor at a time. We highlight the challenges involved in implementing a global sensitivity analysis on the GCAM model, given its large input parameter space and the limited solution convergence space when all input parameters are varied simultaneously. We also implement an illustration GSA on the GCAM model, based on a selection of input parameters that were previously assessed to be most important in the model, relative to climate emissions sensitivity. We conduct GSA on the model using a series of methods and compare the result from the different GSA methods, highlighting the properties of each method. A comprehensive review of the literature related to this work is given in Chapter 2, section 2.3. Additionally we provide in the next section, section 5.1.1, a brief overview of background research that justify the need for this research methodology.
5.1.1 Background

Climate IAMs are policy tools that integrate knowledge from different scientific domains towards modeling the effect of climate change. They provide a consistent framework for policy makers to examine different future scenarios and projections of policy uncertainties in the linkage between human actions and future climate states.

Given that one of the roles of these IAMs is to explore the effect of policy uncertainty on the future state of the climate, considerable need exists for assessing the principal drivers of uncertainty in these models. For this reason several studies have conducted sensitivity and uncertainty analysis on the models (Scott et al., 1999; Traeger, 2013; Draper, 1993; Draper et al., 1987; Webster et al., 2012; Kann and Weyant, 2000; Draper, 1995; Yang, 2011; Cooke, 1991; Santen and Anadon, 2014; Baker and Adu-Bonnah, 2008; Helton and Davis, 2003; Anderson et al., 2014; Golub et al., 2014). Most of these studies, however, have been one factor at a time (OFAT) local sensitivity studies, only exploring a limited parameter space, and not considering all the parameters in the model (Anderson et al., 2014; Saltelli and Annoni, 2010). Saltelli and Annoni (2010) prove that in larger dimensions, when uncertainty is considered in a large number of input parameters, OFAT methods are inadequate for measuring uncertainty as they only explore a very limited fraction of the uncertainty hyperspace. These methods are also not capable of assessing interactions effects between input parameters, nor do they provide a consistent framework for assessing the principal drivers of uncertainty in the models. As such this study focuses on different GSA methodologies that might be implementable in the GCAM model. Global sensitivity analysis analyzes the effect of varying the model parameters simultaneously with a view to apportioning the entire variability in the model to the different parameters or parameter interactions.
5.2 Global Sensitivity Analysis

In this section, we discuss the global sensitivity methodology including the metrics used, the sampling methodology, the research problem and our approach to the problem.

5.2.1 Global Sensitivity Metrics

In this section, we provide a brief overview of the different classes of GSA methodologies, including Variance based GSA, Density based GSA and Distance from CDF based GSA.

**Variance based GSA:** Variance based global sensitivity analysis is a method for estimating how uncertainty in the model’s input parameters affects the model’s output by decomposing the variance of the output parameter into fractions attributable to all the different input parameters (Sobol’, 1993; Wagner, 1995; Saltelli and Annoni, 2010). The method is capable of measuring model interactions across the whole input parameter space while providing a metric that indicates the degree of sensitivity of the model output to each input, based on the model output’s change in variance. They can also handle non-linearities in input output space (Baucells and Borgonovo, 2013).

Intuitively, the method estimates global sensitivity on the model output by calculating the expected reduction in variance of the model output if a model input is fixed (Saltelli et al., 2010). Numerous implementation of this method exist, including Fast Amplitude Sensitivity Testing (FAST) (McRae et al., 1982) and Monte Carlo Integration (Saltelli et al., 2008). The procedure is discussed in detail in Saltelli et al.(2010) and the metric which we will use in our analysis is shown in Equation 37.
Any model can be viewed as a black box and represented by a function $Y = c(X)$, where $X$ is a vector of the $N$ model inputs $X = X_1, X_2, X_3, ..., X_N$ and $Y$ is a scalar representing the model output considered. The inputs are assumed to be random variables that are independently and uniformly distributed over the $N$-dimensional unit hyperspace of the inputs. $X_{-i}$ denotes all inputs aside $X_i$ while $E_{X_i}(Y|X_{-i})$ is the conditional expectation of the model output $Y$, taken over $X_i$ and conditional on all the model inputs aside $X_i$. $Var(Y)$ is the unconditional variance of the model output and $Var_{X_{-i}}[E_{X_i}(Y|X_{-i})]$ is the variance of $E_{X_i}(Y|X_{-i})$ taken over all inputs of the model aside $X_i$.

The metric we will use is, $S_i^T$, the total effect index. It measures the total effect of the input $X_i$ on the model output, including the possible interaction effects with other inputs of any order.

$$S_i^T = \frac{Var(Y) - Var_{X_{-i}}[E_{X_i}(Y|X_{-i})]}{Var(Y)} \quad (37)$$

If $S_i^T$ is greater than $S_j^T$, the model $Y$ is more sensitive to changes in input $X_i$ than $X_j$. This is because the input $X_i$ has more impact on the variance of the model output $Y$, either by itself or through its interaction with other model inputs or groups of more inputs. Note that if $X_i$ is independent of $Y$, then $S_i^T = 0$; however if $Y$ is independent of all variables except $X_i$, then $S_i^T = 1$.

A limitation of this class of GSA is that they assume that the total uncertainty in the output parameters is fully characterized by the first two moments (Baucells and Borgonovo, 2013), which is not the case for skewed distributions. This is because
variance is not an accurate measure of uncertainty, as it considers interactions only up to
the second moment of the output distribution.

**Density based global sensitivity analysis:** Density based GSA introduced in
Borgonovo (2007) tries to overcome the limitations of the variance based methods, by
being moment independent and examining the entire distribution of the model output.
Without restricting itself to any moment of the distribution, the method determines the
proportion of uncertainty attributable to each of the input parameters based on the change
in the density of the output distribution (Anderson et al., 2014).

The importance measure is fully described in Borgonovo (2007) and is
summarized briefly below. As above, we define a model $c(X)$ composed of a set of
uncertain inputs $X_1, X_2, X_3, \ldots, X_N$, where $Y = c(X)$ is the model output and $x =
\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_N$ is a specific realization of $\mathbf{X}$. The following input terminologies are also
defined as follows: $F_\mathbf{X}(x)$ denotes the cumulative distribution function (CDF) of $\mathbf{X}$, i.e.
the joint cumulative distribution of the $X_i$ inputs, $f_\mathbf{X}(x)$ is the corresponding joint density
function of $\mathbf{X}$, and the marginal density of input $x_i$ is given by $f_{X_i}(x_i)$. The following
output related terminology are also defined as follows: $F_Y(y)$ is the CDF of $Y$, $f_Y(y)$ is
the density function of $Y$ and $f_{Y|X_i}(y|X_i)$ is the conditional density of $Y$ given that input
$X_i$ is fixed. Note that $f_{X_i}(x_i)$, the marginal density of $x_i$ is related to the joint density
function of $\mathbf{X}$ as shown in Equation 38.
\[ f_{x_i}(x_i) = \int \ldots \int f_x(x) \prod_{s \neq i} dx_s \quad (38) \]

A measure of the difference between the unconditional, \( f_Y(y) \), and conditional density functions, \( f_{Y|X_i}(y|X_i) \), is denoted as \( s(X_i) \) and is shown in Equation 39 below. The expected difference \( E_{X_i}[s(X_i)] \) is also shown in Equation 40.

\[ s(X_i) = \int |f_y(y) - f_{Y|X_i}(y|X_i)| dy \quad (39) \]

\[ E_{X_i}[s(X_i)] = \int f_{X_i}(x_i) \left[ \int |f_y(y) - f_{Y|X_i}(y|X_i)| dy \right] dx_i \quad (40) \]

Borgonovo then proposes a density sensitivity measure, \( \delta_i \), shown in Equation 41, which is a moment independent sensitivity measure of the effect of model input \( X_i \) on the model output \( Y \). It represents the normalized expected shift in the density of the model output due to the input parameter considered. It is the area between the unconditional output density and the conditional output density relative to a particular input.

\[ \delta_i = \frac{1}{2} E_{X_i}[s(X_i)] \quad (41) \]

The sensitivity measure, Borgonovo Delta, \( \delta_i \), exhibits monotonic transformation invariance (meaning any rescaling of the model parameters does not affect the sensitivity measure). It is also normalized to the unity scale, \( 0 \leq \delta_i \leq 1 \), making it a statistical importance measure (Anderson et al., 2014). The normalization of \( \delta_i \) between 0 and 1 is as a result of multiplying the expected difference of the density functions, \( E_{X_i}[s(X_i)] \), by \( \frac{1}{2} \). This is because the density of the absolute value of the difference of two density functions (similar to \( s(X_i) \)), is triangularly distributed, and by the triangular inequality, the integral of such a density \( s(X_i) \) is upper bounded by the sum of the integrals of the absolute values of the two densities, which sum to 2 (Proof of Property 3, (Borgonovo,
The measure $\delta_i = 0$, when the input $X_i$ is independent of the model output, $Y$, as the conditional density is then equal to the unconditional density. The measure $\delta_i = 1$, when the model output is only dependent on input $X_i$; this is because in this case the conditional and unconditional densities are totally different and non-overlapping. This method is a better indicator of sensitivity measures than variance based GSA since it examines the entire distribution of the output. It is also better than most of the other moment independent GSA methods e.g. the Chun-Han-Tak importance measure (Chun et al., 2000), as it does not make any assumptions on the change in the input parameters (Borgonovo, 2007).

**Distance between CDF’s scale invariant global sensitivity analysis**: Another method for GSA is the invariant probabilistic sensitivity analysis method developed by Bauells and Borgonovo (2013). This GSA method is very similar to the density based GSA method, except the method uses the Kuiper distance (Kuiper, 1960) and the Kolmogorov-Smirnov (Mason and Schuenemeyer, 1983) metrics instead of the Borgonovo Delta, $\delta_i$, for comparing the conditional and unconditional distribution functions. Both metrics are scale and monotonic invariant. Using similar notation as in the other GSA methods, the Kuiper distance, $\beta_{Ku}^i$, and the Kolmogorov-Smirnov, $\beta_{KS}^i$, metrics are shown in Equation 42 and 43 respectively. $\mathcal{F}_{Y|X_i=x_i}$ is the cumulative distribution function of $Y$, conditioned on when the random input $X_i$ is realized as $x_i$.

$$
\beta_{Ku}^i = \mathbb{E}_{x_i} \left[ \Delta P' + \Delta P'' \right] = \int f_{x_i}(x_i) \left[ \sup_y \{ F_{\mathcal{F}_{Y|X_i=x_i}}(y) \} - \sup_y \{ F_{\mathcal{F}_{Y|X_i=x'_i}}(y) \} \right] \, dx_i \tag{42}
$$
\[ \beta_{i}^{KS} = E_{X_i} \left[ \max \left\{ \Delta P', \Delta P^f \right\} \right] = \int f_{X_i} \left( x_i \right) \left[ \sup_{y} \left| F_y \left( y \right) - F_{y|X_i=x_i} \left( y \right) \right| \right] dx_i \] (43)

For the Kuiper discrepancy metric, Equation 42, \( \Delta P^s \) and \( \Delta P^f \) are defined as shown as the maximum shift below and above the unconditional CDF, respectively. The Kuiper discrepancy metric compares two distributions by measuring the expectation over \( X_i \) of the sum of the maximum changes above and below one of the CDF’s. The metric is especially useful for identifying changes in the spreads of the tails of the compared distributions, compared to the Kolmogorov-Smirnov metric, which measures the maximum displacement between the CDF’s, usually at the median values.

The Kolmogorov-Smirnov metric measures the expected maximum absolute value of the difference between the two CDF’s. It is the standard measure of the goodness of fit for comparing distributions. As the metric captures the maximum absolute difference between distributions, it is not a good test for detecting changes in the tail of the distributions; the Kuiper discrepancy excels at this.

These two methods have several advantages over previous GSA methods given that the metrics are scale invariant, are global in perspective, have a low computational burden and do not require additional information about the inputs. We direct the reader to Baucells and Borgonovo (2013) for the detailed discussion of the methodology and the different characteristics of the metrics. We only discuss the low computation advantage given that most large scale IAMs benefit from this characteristic. Traditional estimation of the distance metrics would require \( n \times N^2 \) model evaluations, where \( n \) is the number of input parameters and \( N \) is the sample size of the sensitivity analysis. By post processing the results of a sensitivity analysis and re-sorting the available input samples while maintaining the input output mapping, Plischke (2010) reduced the computational cost to
$N$ which is the number of Monte Carlo simulations needed to properly estimate the input parameters (this is discussed in more detail in Sections 5.2.2 and 5.2.4).

Other GSA methods include expected value of perfect information (EVPI) (Oakley et al., 2010) and correlation based methods, however we do not apply these in the thesis.

The next section discusses the advances in sampling methodologies which enable efficient computation of the GSA metrics by reducing the number of samples needed for estimation convergence.

5.2.2 Sampling Methodologies

To properly characterize uncertainty over the entire input space, a suitable sampling method is required to estimate the high dimensional uncertainty space resulting from the distributions of the inputs considered.

These sampling methods can be broadly classified under traditional Pseudo-Random Monte Carlo methods and Quasi-Random Low Discrepancy methods (Saltelli et al., 2008). We briefly highlight the differences between both sampling methodologies below and direct the reader to their literature (Caflisch, 1998; Sobol, 1998; Saltelli et al., 2008) for a more detailed overview.

**Pseudo-Random Monte Carlo Sampling.** Pseudo-Random Monte Carlo sampling describe methods that sample pseudo-randomly to estimate an underlying distribution function. Given that no truly random systems occur in nature (Sobol, 1998), Pseudo random sampling are the primary means of estimating the distribution without using a deterministic approach such as the Trapezoidal rule. The technique includes
various methods such as Monte Carlo sampling, Stratified sampling, Importance sampling, Sequential Monte Carlo sampling and Mean Field Particle sampling.

For traditional Monte Carlo sampling, the underlying sampling distribution is sampled pseudo-randomly. Based on the central limit theorem, the estimation error of the mean of the sampled distribution obtained from using Monte Carlo sampling approximation, decreases by order $O \left( \frac{1}{\sqrt{N}} \right)$ (Newman and Barkema, 1999) as the number of samples increase, where $N$ is the number of samples.

**Quasi-Random Low Discrepancy sampling.** Quasi-Random Low Discrepancy sampling approximates the multi-dimensional input parameter space based on a low discrepancy sequence estimation of the input distribution space (Sobol, 1998). Unlike Monte Carlo sampling, which generates samples randomly and independently, quasi-random sampling are deterministic, and generate additional samples based on properties of the previously generated samples to enable uniformity and a more even approximation of the sample space (Caflisch, 1998).

A sequence is said to exhibit low discrepancy when all the samples in the sequence are equi-distributed, across the sampling interval. While their fast convergence in estimating a space is a great advantage over pseudo-random sampling, their non-randomness limits their usefulness for purposes such as optimization and simulation. To overcome this lack of randomness, the sequences can be scrambled to mimic randomness, using different techniques, including Reverse Radix and Matousek-Affine-Owen. Figure 17 shows a comparison of samples generated pseudo-randomly, quasi-randomly and quasi-randomly scrambled. Examples of such low discrepancy sequences are the Sobol sequence, Halton sequence, Faure, Niederreiter and the van der Corput sequence. Of all
these, the Sobol sequence has been proven to be superior to all the other low discrepancy sequences (Paskov and Traub, 1995; Jäckel, 2002; Sobol et al., 2011), hence its applicability across a broad range of industries (Glasserman, 2003).

In the next section, we describe the Sobol sequence, the quasi-random Monte Carlo sampling method used in this study.

**Figure 17: Pseudo-Random, Quasi-Random Sobol sequence and Scrambled Quasi-Random Sobol sequence generated samples in respective order.**

**SOBOL Sequence.** The Sobol sequence, introduced by I.M. Sobol (Sobol, 1967), is a quasi-random low discrepancy sequence, which uses base 2 binary expansions
to sequentially estimate uniform partitions of the high-dimensional unit interval representing the input parameter space. The Sobol sequence was created with a view to satisfying three main properties (Saltelli et al., 2008; Lemieux, 2009; Sobol et al., 2011). First, to provide the best uniformity as the number of samples approaches infinity. Second, to ensure a good approximation of the distribution for very small initial sample sizes. Third, it was created to ensure very fast computation of the sequence.

We define a Sobol sequence below by first defining the terms on which the sequence’s definition is based. The following is a review of the formal definition of a Sobol sequence based on Sobol et al.(2011).

- **High-dimension unit hyperspace**: This is a $s$-dimensional unit hyperspace, where $s$ is the number of model inputs considered. We are interested in sampling from this high-dimensional unit hyperspace.

- **Dyadic Interval**: A dyadic interval is any of the $2^n$ intervals formed by dividing the unit interval into $\frac{1}{2^n}$ equal lengths, where $n$ is a non-negative integer. Each unit interval corresponds to an input.

- **Dyadic Box**: A dyadic box is the hyperspace resulting from each collection of the $s$ dyadic intervals. It is denoted as $\Pi$ and is the product of the $s$ dyadic intervals. The High-dimension unit hyperspace is divided into $2^n$ dyadic boxes each with volume $2^{-n}$, with $n = n_1 + n_2 + \cdots + n_s$, where $n_i$ are non-negative integers representing the number of $2^{n_i}$ dyadic intervals each input’s $i$ unit interval is divided into.
- \((t, m, s)\)-net or \(P_t\) net: \(t\) and \(m\) are non-negative integers with \(m > t \geq 0\). Any set of \(2^m\) points in the high-dimension unit hyperspace is defined as a \((t, m, s)\)-net if there are exactly \(2^t\) points of the net in each dyadic box of volume \(\frac{2^t}{2^m}\).

- Dyadic section: A dyadic section of a sequence is defined as any set of points \(q_i\) indexed by \(i\), which satisfy the following condition \((k - 1)2^p < i \leq k2^p\). Where \(q_0, q_1, q_2, \ldots\) is an infinite sequence of points in the high dimension unit space, \(2^p\) is the length of the dyadic section and \(p, k\) are any arbitrary positive integers.

  Section \((q_0, q_1, q_2, q_3)\) of the sequence is a dyadic section of length 4 while \((q_1, q_2, q_3, q_4)\) is not. This is because each point in section \((q_0, q_1, q_2, q_3)\) satisfies the condition \((k - 1)2^p < i \leq k2^p\) at \(k = 1\), while \(q_4\) in the section \((q_1, q_2, q_3, q_4)\) does not satisfy the condition at \(k = 1\).

- Sobol sequence: A sequence of infinite points in the high dimension unit hyperspace is called a Sobol sequence if all the dyadic sections of the length greater \(2^t\) are \((t, m, s)\)-nets. The sequence is a binary expansion estimation of the uncertainty space of the inputs; it ensures even distribution and uniformity. The Sobol sequence is deterministic in nature and this limits its applicability.

  However, methods such as Matousek-Affine-Owen exist for mimicking randomness in the sequence while maintaining low discrepancy. Numerous implementation of this sequence exist, including a modification of the Gray code, based on an implementation by Antonov and Saleev (Antonov and Saleev, 1979).

  The primary purpose for developing quasi-random sampling techniques such as the Sobol sequence is to ensure fast approximation accuracy and low discrepancy of the
space being sampled. Discrepancy is measured in terms of the difference in uniformity between the generated sequence and a perfectly uniformly distributed sample space. Several studies have investigated the discrepancy of these sampling methods. Given that \( N \) is the number of samples and \( s \) is the number of inputs, the upper bound on the discrepancy from using the Sobol sequence can be shown to be of order \( O \left( \frac{(\log N)^s}{N} \right) \) (Lemieux, 2009). The rate of convergence of the Sobol sequence is however found to be closer to order \( O \left( \frac{1}{N} \right) \) in practice (Asmussen and Glynn, 2007), and thus is independent of the number of model inputs. Relative to Monte Carlo sampling, this improvement in convergence of the Sobol sequence, comes at an additional cost in generating the samples. This is because, for subsequent samples of the Sobol sequence to be generated, the properties of the previously generated samples must be taken into account to ensure even sampling of the multi-dimensional space (Saltelli et al., 2008). This is unlike the Monte Carlo sampling method which samples the multi-dimensional input space independently. One limitation of this constraint is that, at current computational capacity, the maximum dimension for which a Sobol sequence can be generated is 1000 inputs (Borgonovo, 2007). This severely limits our GSA of the GCAM as is discussed later in Section 5.3.4.

5.2.3 Problem Scope

As IAMs have become increasingly important for aiding climate policy making, they have also metamorphosed into different variants ranging from purely analytical climate economy models which are representable in a limited set of equations (e.g. DICE
(Nordhaus, 1993)) to very complex bottom up models consisting of a detailed representation of the earth climate system (e.g. GCAM (Clarke et al., 2007)). While a study has carried out GSA on the small scale DICE model (Anderson et al., 2014), no previous study that we know of has attempted a global sensitivity analysis of a large scale IAM as the GCAM model. This is primarily because of the computational cost for the analysis given the very large input parameter space and the detailed nature of the model.

5.2.4 Approach

In this study, we highlight the challenges involved in implementing a global sensitivity analysis of the GCAM model. We show and discuss all the results from the different stages before implementation.

We approach the uncertainty quantification analysis of the GCAM IAM using the global sensitivity methods discussed in section 5.2.1. A series of prior research studies summarized in Anderson et al. (2014) reduced the computational complexity of conducting a GSA, based on sampling using a Sobol sequence, from $Nn^2$ to $N$, where $N$ is the sample size required for a Monte Carlo estimation and $n$ is the number of model parameters considered for global sensitivity.

In the sections below, we discuss the methodology for our study. In section 5.3, we discuss the steps taken towards the GSA study. Given the non-convergence of the GCAM model, when all model parameters are simultaneously varied, we implement an illustration study based on different global sensitivity measures on the most important variables. Section 5.4 discusses the illustration GSA conducted on a few GCAM model inputs.
5.3 Methodology: Global Sensitivity Analysis on all Variables of the GCAM Model

In the following sections, we discuss the steps taken towards the implementation of the global sensitivity analysis in the GCAM model. We start with an overview of the GCAM model structure in section 5.3.1, we then discuss the processing of the model inputs in section 5.3.2, the integration of uncertainty on the input parameters in section 5.3.3, the sampling of the parameter space in section 5.3.4 and the evaluation of the model in section 5.3.5.

5.3.1 GCAM Model Structure

As discussed in earlier chapters, the GCAM model is a detailed large scale integrated assessment energy-economic model composed of interacting agricultural, climate, land use and economic units (Edmonds, Wise et al. 1994, Clarke, Kyle et al. 2008). The model is developed by the Joint Global Change Research Institute of the Pacific Northwest Laboratory in affiliation with the University of Maryland and has been used in most of the recent IPCC assessment reports.

GCAM is a dynamic-recursive global economic model, composed of 14 disaggregated regions of the world which interact through trade and emission policy constraints. It solves in 15 year time steps up to 2095. The model is compiled in C++. To facilitate its transition to a community based climate modeling tool, the inputs to the model are in XML. The implementation of the model consists of a core base configuration file which links the reference parameters values to all the other model components. Figure 21 provides an overview of the hierarchical structure of the base
configuration file in the GCAM model, across the different constituent technologies, regions, resources and other specifications of the model.
Figure 18: Hierarchical tree structure of the base configuration file in the GCAM model. Child nodes are in blue and roots in green.
5.3.2 Model Input Parameters

In this section, we discuss the GCAM model parameters that we analyze, based on the GCAM core-model-input file. When the base configuration file is parsed for all input parameters, the total number of strict numeric input parameters in the model is 47,125. As shown in the tree structure of the model in Figure 21, most of the input parameter nodes are similar in structure but differ in that they describe different sections of the model.

Given that there is no uncertainty on past data, we exclude inputs before 2020 from our sensitivity analysis. The total number of remaining parameters is 29,122.

5.3.3 Input Parameter Distribution

To integrate uncertainty into the input parameters, we assume a triangular distribution about each of the input parameters to indicate a lack of knowledge about the true state of the considered parameters. $a, b, c$ are the lower limit, upper limit and mode of the triangular distribution, and are defined as $a = \frac{2\times\text{GCAM Value}}{3}$, $b = \text{GCAM Value}$ and $c = \frac{4\times\text{GCAM Value}}{3}$ respectively. We assume each of the parameters is symmetrically distributed with the mean $\frac{a+b+c}{3}$ and mode $c$ coinciding at the current GCAM value.

5.3.4 Sampling Methodology

Our methodology and the improvement in computational complexity depends on sampling from a low discrepancy sequence to get a very even approximation of the parameter distribution sample space. By low-discrepancy, we mean a sequence that is uniform equi-distributed over the input parameter hyperspace. We generate a Sobol
Sequence of $N$ by $n$ samples as required by Equation 45, where $N = 10,000$ is the required number of the Monte Carlo samples required (Plischke et al., 2013) and $n = 29,122$ is the number of model input parameters (The MathWorks, 2012). That is, we generate 10,000 samples with 29,122 values in each sample.

Note, the Sobol process approximates the parameter space more closely by using a base of two to form successive finer uniform partitions based on a unit interval (Sobol’, 1993). We transform the generated, unit interval uniformly distributed, Sobol set to the triangular distribution using the Monte Carlo inversion theorem (Owen, 2013). This makes use of the fact that all CDF’s of any distribution function is cumulative over the unit interval. Hence, we can easily transform the uniform distribution to any desired distribution.

The cumulative distribution function of the triangular distribution, $D(x)$, shown in Equation 44, and the inverse transform, $F^{-1}(u)$, is given in Equation 45. We transform each of the Sobol sequence generated using Equation 45 to obtain a triangularly distributed Sobol sequence. We then multiply the actual GCAM values with the generated values to obtain a triangular distribution.

---

6 Due to difficulty in maintaining an uniform representation of the sample space, the current maximum possible Sobol set that can be generated is 1,000. We therefore approximate the generation of an evenly dispersed set of 10,000 samples by generating 10 sets of 1,000 by 29,122 Sobol sets. We acknowledge a loss of accuracy here.
\[
D(x) = \begin{cases}
\frac{(x-a)^2}{(b-a)(c-a)} & \text{for } a \leq x \leq c \\
1-\frac{(b-x)^2}{(b-a)(c-a)} & \text{for } c < x \leq b
\end{cases}
\] (44)

\[
F^{-1}(u) = \begin{cases}
\frac{a+\sqrt{(b-a)(c-a)u}}{b-a} & \text{for } a \leq u \leq \frac{c-a}{b-a} \\
\frac{b-\sqrt{(b-a)(c-a)(1-u)}}{b-a} & \text{for } \frac{c-a}{b-a} < u \leq b
\end{cases}
\] (45)

5.3.5 Model Evaluation

Once the input parameters have been transformed as discussed as above, we parse the inputs into the GCAM model and evaluated the model under a no policy reference case scenario. We find however that the model does not converge, indicating that varying all the input parameters at once leads to infeasible relationships between parameters which do not satisfy certain preset bounds in the model solver algorithm.

This indicates one of the major challenges of conducting GSA of detailed bottom up IAMs, as these models have complex linkages and relationships, giving them only a very limited solution range. Given the non-convergence of the model, we therefore carry out an illustration global sensitivity analysis based on a set of variables that were previously assessed to be the most significant drivers of uncertainty. This study is for illustration purposes only, as the Evergreen computing cluster was recently retired, restricting large scale evaluation of the GCAM model. We discuss the methodology for the study in section 5.4 below.
5.4 Methodology: Illustrative Global Sensitivity Analysis study on a Limited Set of Variables in the GCAM Model

In this section, we discuss our methodology towards conducting an illustrative GSA analysis on a set of input parameters in the GCAM model. Our analyses rely on a previous study by Scott et al. (1999). This paper conducted an OFAT sensitivity analysis of the MiniCAM model, with a view to determining the inputs that are the primary drivers of uncertainty in relation to model outputs, including the Atmospheric Carbon Emissions, Year 2100 Carbon Concentration, Temperature rise and Market damages. The study finds that, for most of the GCAM model outputs, the three main drivers of uncertainty in the model are the Income Elasticity of Demand, Labor Productivity of Demand and the Autonomous Energy Intensity Improvement Index (AEII) for different regions and sectors.

We reprise a smaller version of this OFAT study to examine if the ranking of the input parameters is still as assessed in the Scott et al. (1999) study, given that the MiniCAM and GCAM models are slightly different. We then conduct a GSA analysis on the three selected input parameters relative to the radiative forcing output of the GCAM model, assuming a no policy scenario. While the Scott et al. (1999) study evaluated uncertainty over several of the model’s outputs, we focus on the radiative forcing output as it is a good indicator of climate sensitivity (PALAESENS Project Members, 2012). The GSA is assessed via the metrics discussed in Section 5.2 and the results compared to those of the OFAT study to determine if there is sufficient interactions between the input parameters to cause a change their ranking in terms of their effect on model variability. We discuss the GSA and OFAT methodologies in detail, in the sections below.
5.4.1 Input Parameters Considered

Based on the previous study by Scott et al.(1999), we considered the three input parameters that were shown to be the most significant drivers of variability in the atmospheric emission concentration and the temperature rise. These are the Labor Productivity Growth Multiplier, the Income Elasticity of Demand and the Autonomous Energy Intensity Improvement Index (AEII). We integrate uncertainty into these inputs as their 2020 values, but maintain the year on year trend as in the current GCAM model. A brief discussion of the input parameters is given below.

*Labor Productivity Growth Multiplier:* This indicates the growth in gross domestic product per capita. From the Scott et al.(1999), this was the second most important driver of variability in the climate emissions. This is intuitive given that, when labor productivity increases, consumption increases, leading to higher utilization of resources and hence more impact on the climate.

Equation 46 (Clarke et al., 2007) shows the relationship between the labor productivity growth multiplier $Pro_{1,m}$ and GDP. In the equation, $GDP_{index_{1,m-1}}$ is the GDP value normalized against the base year GDP, $Pro_{1,m}$ is the labor productivity growth multiplier on which uncertainty is evaluated, $Nstep$ is the model resolution between periods and $LBFindex_{1,m}$ is the ratio of the current labor force population relative to the previous period’s labor force population. The current value of the labor productivity growth multiplier for the year 2020 Industry sector is given as 0.01602.

$$GDP_{index_{1,m}} = GDP_{index_{1,m-1}} \ast (1 + Pro_{1,m})^{Nstep} \ast LBFindex_{1,m}$$ (46)
**Income Elasticity of Demand:** This was assessed to be the primary driver of variability in climate emissions and temperature rise in the Scott et al. (1999) study. It represents the sensitivity of consumer demand to changes in income of the consumer.

Equation 47 shows the linkage between the income elasticity of demand and the demand for energy (Clarke et al., 2007), where $P$ is the energy service price, $X$ is the GDP index, $b$ is a scale parameter, $r_{pk}$ is the energy price elasticity of demand and $r_{xk}$ is the income elasticity of demand for the industry sector.

\[
Demand = b \times P^{r_{pk}} \times X^{r_{xk}}
\]  

(47)

The current year 2020 GCAM value for the US Industry sector is -0.155.

**Autonomous Energy Intensity Improvement Index (AEII):** This is the primary indicator of technological change in the GCAM model. It is used to account for the diffusion of technical change and reflects the change in the energy intensity to gross domestic product ratio, holding energy prices constant (IPCC4, 2007). A high AEII indicates a lower cost of adopting clean energy technology hence low cost of climate mitigation. Equation 48 represents the integration of the AEII into the GCAM model (Clarke et al., 2007), where $Tech_{j,k,m}$ and $Tech_{j,k,m-1}$ represents the level of technological progress for period $m$ and $m - 1$ respectively, while $T_{k,1,m}$ is the Autonomous Energy Intensity Improvement Index (AEII) for period $m$.

\[
Tech_{j,k,m} = Tech_{j,k,m-1} \times (1 + T_{k,1,m})^{Naep}
\]  

(48)

The current year 2020 GCAM value for the US industry sector is 0.0034.
We implemented a GSA and an OFAT on the GCAM for the three input parameters listed above, restricting ourselves to the US regional values of these parameters for the Industry sub-sector. This is necessary given that the large number of input parameters that exist for each region and subsector as shown in Figure 21.

5.4.2 Output Parameter Considered

We considered the effect of uncertainty in the input parameters relative to the radiative forcing of the earth in 2100. It is defined by the IPCC as the balance of the incoming insolation absorbed by the earth’s atmosphere and the outgoing energy radiated back (IPCC4, 2007). It is measured in Watts/M². It has been used as a summary statistic of the effect of climate emissions and temperature rise in a lot of the recent literature. Given the nature of the GCAM model, at no additional memory expense, we obtain a source file containing different possible outputs of the model and can therefore conduct similar analysis on other output variables.

5.4.3 Input Parameter Distribution and Sampling Methodology

We assume a normal distribution over each of the input parameters considered, with the mean and standard deviation assumed to be equal to each other, and equal to the current GCAM value for the parameter. Given the cost of model evaluation in the absence of a computing cluster, we generate a Sobol sequence of 128 samples.

We transform the uniformly distributed Sobol set to the normal distribution using the Monte Carlo inversion theorem (Owen, 2013) using the CDF transform. As the CDF of the normal distribution is 
\[ D(x) = \frac{1}{2} \left[ 1 + erf \left( \frac{x - \mu}{\sigma \sqrt{2}} \right) \right] \] 
where \( \mu \) and \( \sigma \) are the mean and

122
standard deviation of the normal distribution respectively and are both equal to the current input parameter value, the Monte Carlo inversion theorem gives the inverse CDF as shown in Equation 49, where $F^{-1}(u)$ is the inverse CDF of the normal distribution and $u$ is the cumulative probability of the normal distribution of the random variable $x$.

$$F^{-1}(u) = \mu + \sigma \sqrt{2} \left[ \text{erfinv}(2u - 1) \right]$$

(49)

We transform the Sobol sequence generated using this transform to obtain a normally unit interval distributed sample. We then multiply the GCAM values with this normally unit interval distributed sample to obtain a normal distribution of the current GCAM values. We then evaluate the GCAM model with each set of samples generated and obtain the radiative forcing output for each model run.

Here we discuss the limitations due to the small sample size used. There is no work in the review of the literature that estimated the number of samples necessary for a multi-dimensional sampling of the input parameter space generated by a Sobol sequence. In most of the literature (Yang, 2011; Burhenne et al., 2011; Plischke et al., 2013; Anderson et al., 2014; Borgonovo et al., 2014), the studies evaluate the GSA sensitivity metrics over a wide range of samples sizes to examine the effect of sample size on the validity of the results obtained. The papers reviewed found that the GSA metrics appear to be robust to low sample sizes, in the sense that they provide the same ranking of which parameters are most important. For example, Plischke (2013) found that as few as 512 samples produced the same ranking as 65,536 samples, albeit with the ranking at 65,536
samples showing clearer separations between the ranks of the inputs. Their work was based on a model consisting of 872 model inputs. Similarly, for a 4 input model, Burhenne (2011) conducted mean squared error to sample size analysis (upto to 512 samples) and found at a sample size of 64, the mean squared error between the sample and the actual distribution was zero. The authors go on to recommend Sobol sequence sampling as the appropriate sampling methodology when the cost of evaluation prohibits large sample sizes. For our 3 input model, we chose a sample size of 128 due to the difficulty in evaluating the model. We note, that the distributions of the 3 inputs, using the 128 samples appear to be somewhat normally distributed (Figure 22) but with more weight in the tails.

From our discussion in Section 5.2.2, we note that even at such a small sample size of \( N = 128 \), the Sobol sequence has less initial estimation error and also faster convergence than a pseudo-random normally distributed sample. Evaluated at a sample size of \( N = 128 \) and an input dimension of \( s = 3 \), the upper bound on the Sobol sequence error is less than that of a pseudo-random normal sample \( O \left( \frac{(\log N)^s}{\sqrt{N}} \right) < \)
\[
O \left( \frac{1}{\sqrt{N}} \right)
\]
i.e. \( 0.0731 << 0.0884 \) and the usual bound on the Sobol sequence error in practice is much less than that of the pseudo-random normal error \( O \left( \frac{1}{\sqrt{N}} \right) << \)
\[
O \left( \frac{1}{\sqrt{N}} \right)
\]
i.e. \( 0.0078 << 0.0884 \). From Figure 20, we also note that our normally transformed Sobol sequence gives a better estimation of the multi-dimension input space relative to if the sample were pseudo-randomly generated in the sense that the Sobol sequence show better clustering in the middle and also has better approximation of the tails of the distributions, compared to the pseudo-random normal sequence. The figure
compares the 3 dimensional space of the normally transformed Sobol sequence with a generated pseudo-random normal distributed sample with the same mean and standard deviation.

Figure 19: Resulting Model Input and Output Distribution.
Figure 20: Three Dimensional Plot of the Input Parameter Space Distribution.  
*The Blue Figure Shows The Actual Normally Transformed Sobol Sequence Used For Our Study. The Red Figure Shows A Hypothetical Pseudo-Random Normal Distribution Of The Same Sample Size.*

5.4.4 Global Sensitivity Analysis Methodology

We implement a GSA analysis using each of the metrics discussed in section 5.2; the variance based GSA, density based GSA and the distance from CDF GSA methods.

5.4.5. One Factor at a Time (OFAT) Sensitivity Analysis

In addition to a selective GSA analysis, we also implement an OFAT analysis as discussed in Scott et al.(1999). We do this by examining the effect of a 50% increase in each of the three inputs on the 2100 cumulative radiative forcing output of the model. The absolute percentage change in the radiative forcing output of the model, when we go from the base GCAM values to the higher values, is a measure of the sensitivity of the model to each of the inputs.

We note that this study differs from the Scott et al.(1999) study in three ways. First, the Scott et al.(1999) study was conducted on the MiniCAM model, in contrast to
our study which is conducted on the GCAM model. Second, the Scott et al. (1999) study conducted an OFAT analysis on all parameters in the MiniCAM model, but this study restricts its scope to three parameters only in the Industry Sector of the USA region. Finally, the Scott et al. (1999) study evaluated the sensitivity based on different outputs of the model (Year 2100 carbon concentration, Temperature rise, Market damages). As all the outputs showed the same ordinal ranking to sensitivity in each of the inputs, we focus on the Radiative Forcing output to summarize the impact of sensitivity in the inputs.

We discuss the results of both the OFAT analysis and the GSA analysis in the next section.

5.5 Results

5.5.1 Global Sensitivity Analysis of the GCAM model

This section discusses results from an attempt to conduct GSA on the GCAM; while this was not ultimately possible, we discuss our findings from the different stages of the analysis.

The goal of this study was to identify the principal drivers of variability in the model with a view to helping integrated assessment modelers understand how intra-model interactions affect the GCAM model output. To do this we conduct a GSA over all the model parameters in GCAM, by making changes to the base configuration file. We however could not get the model to solve or converge possibly due to narrow solution range which is not satisfied when all model inputs are change simultaneously. The implementation codes and base files are provided in this link.
One possible future recommendation for approaching global sensitivity problems in models of this scale is to exploit the hierarchical structure of the GCAM model by conducting GSA on a much smaller subset of the model e.g. a region or sub-sector in a region. As an example, GSA can be carried out on only the USA region of the model. As the maximum number of inputs to be sampled using a Sobol sequence is 1000, the number of inputs considered in the desired region would have to be the 1000 most important model inputs, identified using an OFAT study.

Another possibility is to examine systemic ways of reducing the number of parameters in the GSA. This can be done by binning the input parameters into different categories depending on their role in the model (Campolongo et al., 2007). As an example, all population related inputs could be classified as a single input parameter, as well as other classes of parameters, such as efficiency improvement-related parameters and others.

Another challenge is the non-convergence of the GCAM model due to the narrow solution range of the model when all the input parameters are simultaneously varied. One possible recommendation to overcome this is to provide different uncertainty probability distributions and variability ranges on different model inputs. While the use of different uncertainty probability distributions and variability range on different inputs is itself subjective, this might help with convergence of the model as the variability range of each input parameter can be personalized depending on how variable the parameter is in reality. This is because there are different states of knowledge on the different model inputs. This ensures that the combination of input parameters in the model is not out of the solution convergence range of the model.
5.5.2 Illustration Global Sensitivity Analysis on a Selection of Variables

As discussed in Section 5.4, we conduct both an illustration GSA analysis and an OFAT analysis on three of the previously assessed most important variables from an OFAT study in Scott et al. (1999). From the OFAT study we observed that the Labor Productivity Growth and income Elasticity of Demand have a positive relationship with Radiative Forcing while the AEII has a negative relationship with the Radiative Forcing output. This is expected as technology advancement will reduce the cost of carbon abatement while increase in productivity will lead to increase in consumption, hence more impact on the climate.

We discuss the results from both the OFAT and GSA analyses below.

Figure 21 shows the result from the OFAT and GSA studies. To enable ease in comparing the ranking of the different metrics, each of the input metrics displayed is normalized to the 0-1 scale by dividing with respect to the sum of the three inputs for each metric. Each metric shown in Figure 21 can therefore be interpreted as a proportion of the uncertainty in the Radiative Forcing output due to the particular input using the specific metric, since the sum of the GSA metrics over all the parameters considered is not necessarily 1.

We find that all the GSA methods and the OFAT study result in the same ranking of input parameters in terms of their effect on the Radiative forcing output of the model. Of these three inputs considered, the Labor Productivity Growth Multiplier is the most significant driver of uncertainty in the Radiative Forcing output of the GCAM model, followed by the AEII and the Income Elasticity of Demand. Our results are contrary to the Scott et al. (1999) study, which find Income Elasticity of Demand is the most
significant driver of uncertainty in the model followed by Labor Productivity Growth and AEII respectively.

Though several studies (e.g. (Saltelli and Annoni, 2010; Saltelli et al., 2008)) have shown that OFAT analysis is not an accurate measure of uncertainty in models, we find that we obtain the same ranking of the input parameters using the two methodologies. This is likely due to the illustrative nature and small scale of the global sensitivity study. Specifically, this is because the number of input parameters considered is only three, which is small relative to the large number of input parameters in the GCAM model. This likely results in insufficient interaction between the inputs to lead to a change in the ranking of the inputs, as the additional contribution of the higher order interactions is not enough to offset the first order contributions as assessed and shown in the OFAT study.

To reiterate, the OFAT study only examines first order effects, and the Variance based GSA also has a limited assumption that variance is a total description of uncertainty. Also, while easier to compute, both CDF-based metrics have their limitations, as the Kolmogorov-Smirnov metric does not excel in measuring tail spreads and the Kuiper distance focuses overly on tail spreads.

From Figure 21, we observe that, with the exception of only the Kuiper distance, when higher interactions are examined with a GSA analysis, the contributions of the Income Elasticity of Demand and the AEII increases significantly. This shows that Income Elasticity of Demand and AEII play a more significant role in driving uncertainty in the GCAM model than was shown using the OFAT analysis. The result from the Kuiper discrepancy metric is slightly different from the other GSA metrics due to the small number of samples used in estimating the distributions. As discussed in section 5.2,
the Kuiper discrepancy excels at finding differences in the tail of compared distributions. Given that we use a small sample size, the tail of the compared distributions is distorted, as can be seen in Figure 19: the input distributions are observed to be heavy tailed compared to a typical normal distribution. This leads the Kuiper discrepancy metric to produce a slightly different result compared to the Kolmogorov-Smirnov test which measures the changes at the mean of the normal CDF’s compared.

By comparing the Variance based GSA to the Density based GSA and the Kolmogorov-Smirnov metric, we observe that higher interactions between the input parameters play only a slight role, given the similarity between the three metrics. We do note that the income Elasticity of Demand and the AEII play a slightly more significant role in the Density Based GSA, relative to the Variance based GSA.

![Figure 21: Normalized OFAT and Global Sensitivity Analysis GSA analysis metrics. Based on their effect on the Radiative Forcing Output of the GCAM model.](image-url)
5.6 Conclusion

In this study we highlight the challenges involved in conducting a global sensitivity analysis of the GCAM model. We note the non-solvability of the model when all input parameters are varied simultaneously, possibly due to limited bounds between the model inputs given. We also note that, given current computation resources, such large scale global sensitivity studies are not possible as the number of variables involved vastly exceeds the current maximum for concurrent generation of equi-distributed quasi random variables.

We also conducted an illustration global sensitivity analysis based on some input parameters that were previously assessed as the most important model variables, from a one factor at a time (OFAT) sensitivity analysis on the GCAM model. We find that we obtain the same ranking of the most important input parameters using both the GSA and OFAT analysis. This is likely due to the illustrative small scale of the GSA study. We however still observe the value of using a GSA analysis, as we see an increase in magnitude of the contribution of the Income Elasticity of Demand compared to the other inputs, when interaction effects are considered using the GSA. We also find that the parameter importance ranking for both the OFAT and GSA studies is different from that obtained from the Scott et al. (1999) study; this is likely because the Scott et al. (1999) study was based on the MiniCAM model, which is the immediate predecessor to the GCAM model.
CHAPTER 6

CONCLUSION AND FUTURE RESEARCH RECOMMENDATIONS

This thesis examined three topics centered on the role of low carbon energy technologies in near term energy policy, in the face of climate change. The first part of the thesis focuses on a large scale scenario analysis of clean energy technology options. The second part of the thesis addresses the optimal allocation of R&D funds to clean energy technologies and the third part focuses on global sensitivity analysis of one of the energy models used in our study. In the sections below we summarize the results from each of these 3 parts of the thesis and we also discuss possible future research recommendations.

From Part 1, the large scale scenario analysis of low carbon energy options, we approach the near term energy policy problem from a global scenario modeling perspective. Using an integrated energy assessment model, GCAM, all possible energy scenarios that can result from the combination of our R&D technology targets are modeled. We integrate these scenarios into a more succinct integrated assessment model, DICE and evaluate the utility of having these scenarios in terms of their economic impact. We conduct scenario analysis on this set of energy scenarios. Additionally, we conduct probabilistic portfolio analysis based on certain portfolios which describe the extremities of the energy portfolio options. Using stochastic dominance techniques, we compare the different portfolios under different climate damage uncertainty cases.

One of our results from the first part of the thesis is that we show that R&D in CCS and Nuclear technologies is important, due to their low cost of carbon abatement. We also find that the value of technological advancement in technologies is a function of
the uncertainty in climate damages with CCS R&D development increasing in value and Solar R&D development reducing in value, as the uncertainty in climate damages increases. We also show that Bio-Electricity and CCS are significant energy complements while most of the other technology pairs are substitutes. Other findings from this part of the thesis include that there is a need for directed portfolio analysis as smaller funding portfolios can stochastically dominate much larger portfolios; and that the dominance between the portfolios further increases with higher climate damage uncertainty. For example, an economical portfolio with only Bio-Electricity and CCS stochastically dominates an expensive portfolio of all renewables and the degree of dominance between these portfolios increases with climate damage uncertainty. We also show that certain portfolios are robust to climate damage uncertainty while other portfolios are best under a specific climate uncertainty case. For example, a medium funding portfolio, of all the technologies, is robust to climate damage uncertainty and portfolios with Bio-Electricity and CCS funding are best under the high climate damage uncertainty case.

For the second part of this thesis, R&D allocation in a large portfolio of 6 clean energy technologies, we approach the near energy policy problem from an optimization perspective. We model the problem sequentially in stages (scenario modeling, scenario evaluation and R&D optimization), and use a genetic based optimization approach to obtain the best R&D portfolio for the different climate uncertainty cases. We also analyze how the resulting portfolio can be decomposed optimally, without re-optimization.

One of our findings from this part of the thesis is that there exists significant returns to scale in R&D investment indicating that there is a strong need for R&D
portfolio optimization. We observe that the resulting optimized portfolio varies significantly with climate damage uncertainty and that as climate damage uncertainty increases so do the investment in R&D of Bio-Fuels and CCS at the expense of investment in R&D of Solar, Batteries for Electric Transportation and Bio-Electricity. We notice however that investments in R&D of Nuclear energy are robust, to climate damage uncertainty, as they lead to significant reduction in the cost of carbon abatement but they do so at very high R&D costs. Another major result from this part of the thesis is that we show that portfolio problems can be decomposed optimally and that the solutions to the decomposed portfolio can be obtained, without optimization, provided no interactions exist between the constituent technologies. This is valuable given that there is a constant flux of promising new technologies and that there might exist a need to replace a technology in an already optimized portfolio with another promising technology.

One possible future research recommendation on the R&D portfolio allocation problem is the evaluation of the optimal portfolio using exact optimization techniques, such as Branch and Bound. Another recommendation is the exploration of decomposition techniques to exactly quantify the change in the portfolio, when significant interactions exists between the constituting technologies. Another very significant future research recommendation is the exploration of dynamic constrained portfolio optimization, for incremental portfolio problems, using the decomposition methodology described in this chapter of the thesis. For example, can the optimal portfolio for a computationally intractable 10 technology problem be estimated based on solutions to the tractable 9 technology sub problems? We believe that there is a need for further exploration of this
concept given the limited similarity between the portfolio decomposition and the portfolio extension problems.

For the third part of the thesis, global sensitivity analysis of the global change assessment model, we examined the impact of input parameter variability on the GCAM model through a global sensitivity analysis. Based on limitations in implementing this methodology on the model, we conduct a limited global sensitivity analysis on the GCAM integrated assessment model, based on a set of parameters found to have the most impact on the predecessor of the GCAM model. We use the variance based, density based and distribution based global sensitivity metrics to measure and rank the impact of these input parameters on the model.

One of our findings from this part of the thesis is that we observe that the narrow model solution convergence range of the GCAM model and the limitations of the current best available computational capabilities hinder the implementation of global sensitivity analysis on such a large model as GCAM. We also find from the limited global sensitivity analysis study that when the number of input parameters considered is low, global sensitivity analysis can lead to the same importance ranking of the parameters, as with a one-factor-at-a-time sensitivity analysis. This is because in the case that the inputs considered are a minute subset of all possible inputs, the higher order interactions between input parameters might not significant enough to overcome the individual contributions of the input parameters.

A possible future research recommendation for this part of the thesis is to explore different means for systematic reduction of the number of input models parameters considered. Such methods can include screening of the inputs based on a one-factor-at-a-
time analysis or clustering input parameters into a much smaller set based on their function in the model. Another direction is to exploit the hierarchical nature of the model to classify the inputs into a smaller more manageable subset.

Another potential research recommendation to overcome the non-convergence of the model is to explore tailored probabilistic distributions for each inputs based on the assessed uncertainty of these inputs. This can include the use of different distributions for each inputs as well as different uncertainty ranges.

Given the large number of inputs considered, another research recommendation, is the use of cloud-based distributed computing and the big data tools such as Hadoop and Pig, for data processing. The deployment of these tools can aid the research process as most of the generated files are several gigabytes. This will help address the severe lack of literature on the application of global sensitivity analysis to systems that consist of large number of interconnected models.

Another potential future research recommendation is to examine the propagation of uncertainty in multi-model ensembles as given in our R&D portfolio ensemble of models. Because several modeled systems, consist of a number of interconnected models, with many of such model inputs depending on the outputs of other models, the assessment of global uncertainty in such conjoined systems is critical.
APPENDIX A

R&D FUNDED TECHNOLOGIES, SUB-TECHNOLOGIES, FUNDING LEVELS, SUCCESS PROBABILITIES AND R&D ENDPOINTS.

Table A.1 gives the R&D funding, elicited success probabilities and the names of the R&D targets (endpoints) for each of the sub-technologies.

**Table A.1: R&D funding, success probabilities and cost endpoints for all technologies and sub-technologies considered.**

*Where a Technology endpoint is a specific energy cost or performance parameter after R&D development in a sub-technology project.*

<table>
<thead>
<tr>
<th>Technology</th>
<th>Sub Technology</th>
<th>Funding (NPV in $M)</th>
<th>Success Probability</th>
<th>Cost Endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS</td>
<td>Pre Combustion</td>
<td>386</td>
<td>23%</td>
<td>CCS-Pre</td>
</tr>
<tr>
<td></td>
<td></td>
<td>154</td>
<td>11%</td>
<td>CCS-Pre</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39</td>
<td>5%</td>
<td>CCS-Pre</td>
</tr>
<tr>
<td></td>
<td>Chemical Looping</td>
<td>56</td>
<td>16%</td>
<td>CCS-Chem</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38</td>
<td>14%</td>
<td>CCS-Chem</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>2%</td>
<td>CCS-Chem</td>
</tr>
<tr>
<td></td>
<td>Post Combustion</td>
<td>519</td>
<td>93%</td>
<td>CCS-Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>224</td>
<td>86%</td>
<td>CCS-Post</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52</td>
<td>68%</td>
<td>CCS-Post</td>
</tr>
<tr>
<td>NUCLEAR</td>
<td>LWR</td>
<td>346</td>
<td>60%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>260</td>
<td>34%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>173</td>
<td>21%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td>HTR</td>
<td>3089</td>
<td>25%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1544</td>
<td>14%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11%</td>
<td>$1500/KW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>772</td>
<td>1%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2%</td>
<td>$1500/KW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fast Reactors</td>
<td>15443</td>
<td>16%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44%</td>
<td>$1500/KW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4633</td>
<td>1%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32%</td>
<td>$1500/KW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1158</td>
<td>0%</td>
<td>$1000/KW</td>
</tr>
<tr>
<td>Source</td>
<td>Type</td>
<td>Quantity</td>
<td>Efficiency</td>
<td>Cost/Unit</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------</td>
<td>----------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SOLAR</td>
<td>Organic</td>
<td>830</td>
<td>3%</td>
<td>3 cents/KWH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14%</td>
<td>5 cents/KWH</td>
</tr>
<tr>
<td></td>
<td>Inorganic</td>
<td>77</td>
<td>9%</td>
<td>5 cents/KWH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39</td>
<td>15%</td>
<td>3 cents/KWH</td>
</tr>
<tr>
<td></td>
<td>3rd Gen</td>
<td>386</td>
<td>9%</td>
<td>3 cents/KWH</td>
</tr>
<tr>
<td>BIO-FUELS</td>
<td>STP1</td>
<td>1293</td>
<td>76%</td>
<td>BF-stp1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>791</td>
<td>52%</td>
<td>BF-stp1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>97</td>
<td>25%</td>
<td>BF-stp1</td>
</tr>
<tr>
<td></td>
<td>STP2</td>
<td>1293</td>
<td>57%</td>
<td>BF-stp2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>791</td>
<td>33%</td>
<td>BF-stp2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>97</td>
<td>17%</td>
<td>BF-stp2</td>
</tr>
<tr>
<td></td>
<td>HYDROLYSIS</td>
<td>2471</td>
<td>48%</td>
<td>BF-hy-High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10%</td>
<td>BF-hy-Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>77</td>
<td>18%</td>
<td>BF-hy-High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>8%</td>
<td>BF-hy-Low</td>
</tr>
<tr>
<td></td>
<td>GASIFICATION</td>
<td>1413</td>
<td>38%</td>
<td>BF-gas-High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13%</td>
<td>BF-gas-Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>85</td>
<td>7%</td>
<td>BF-gas-High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2%</td>
<td>BF-gas-Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
<td>6%</td>
<td>BF-gas-High</td>
</tr>
<tr>
<td>BIO-ELECTRICITY</td>
<td>STEAM</td>
<td>216</td>
<td>39%</td>
<td>BE-stm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>87</td>
<td>21%</td>
<td>BE-stm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>3%</td>
<td>BE-stm</td>
</tr>
<tr>
<td></td>
<td>GASIFICATION</td>
<td>772</td>
<td>49%</td>
<td>BE-gas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>241</td>
<td>19%</td>
<td>BE-gas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>77</td>
<td>5%</td>
<td>BE-gas</td>
</tr>
<tr>
<td>BATTERIES</td>
<td>LITHIUM METAL</td>
<td>309</td>
<td>9%</td>
<td>Lit. metal-High</td>
</tr>
<tr>
<td>FOR ELEC. TRANS.</td>
<td></td>
<td></td>
<td>19%</td>
<td>Lit. metal-Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>77</td>
<td>3%</td>
<td>Lit. metal-High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8%</td>
<td>Lit. metal-Low</td>
</tr>
<tr>
<td></td>
<td>LITHIUM ION</td>
<td>541</td>
<td>32%</td>
<td>Lit. Ion - High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46%</td>
<td>Lit. Ion - Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>232</td>
<td>15%</td>
<td>Lit. Ion - High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>28%</td>
<td>Lit. Ion - Low</td>
</tr>
</tbody>
</table>
APPENDIX B

INTEGRATING TECHNICAL CHANGE

\[ \bar{c}_D(\mu_t) = \pi_t, \theta_t, \mu_t^{\beta_2} \]  

(B.1)

\[ \frac{d(\bar{c}_D(\mu_t))}{d\mu_t} = \pi_t, \theta_t, \theta_2, \mu_t^{\beta_2} \]  

(B.2)

\[ \frac{d(c_D(\mu_t))}{d\mu_t} = \frac{d(\bar{c}_D(\mu_t))}{d\mu_t} \left[ 1 - (\alpha_0 + \alpha_t, \mu_t + \alpha_2, \mu_t^2 + \alpha_3, \mu_t^3) \right] \]  

(B.3)

\[ c_D(\mu_t) = \int \pi_t, \theta_t, \theta_2, \mu_t^{\beta_2} \left[ 1 - (\alpha_0 + \alpha_t, \mu_t + \alpha_2, \mu_t^2 + \alpha_3, \mu_t^3) \right] \]  

(B.4)

\[ c_D(\mu_t) = \pi_t, \theta_t, \theta_2, \mu_t^{\beta_2} \left[ \frac{1 - \alpha_0}{\theta_2} - \frac{\mu_t \alpha_1}{1 + \theta_2} - \frac{\mu_t^2 \alpha_2}{2 + \theta_2} - \frac{\mu_t^3 \alpha_3}{3 + \theta_2} \right] \]  

(B.5)
APPENDIX C

ENERGY SCENARIOS AND R&D OUTCOMES

Table C.1 gives the number of energy scenarios and sub-technological outcomes possible from our portfolio of technologies. The third column and fourth columns give per sub-technology the number of success Technology endpoints possible and the total number of Technology endpoints possible if failure of the R&D process is considered, respectively. Columns 5 and 6 of the table give these endpoints respectively at the technology level. The number of possible R&D outcomes is 2,239,488 while the number of scenarios is 3,780 as a function of the sub-technology endpoints and technology endpoints respectively.

Table C.1: Scenarios and Outcomes Definition and Calculation.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Sub-Technology</th>
<th>Success endpoints per Sub-Tech</th>
<th>Endpoints per Sub-Tech</th>
<th>Success endpoints per Tech</th>
<th>Endpoints per Tech</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS</td>
<td>Pre-Combustion</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Chemical Loop</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post Looping</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuclear</td>
<td>LWR</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HTR</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>FR</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solar</td>
<td>Organic</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Inorganic</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3rd Gen</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bio-Electricity</td>
<td>Steam</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Gasify</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bio-Fuels</td>
<td>Thermal Proc. 1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thermal Proc. 2</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hydrolysis</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Gasification</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batteries for Elec. Trans.</td>
<td>Lithium Metal</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Lithium Ion</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Technologies</td>
<td>17 Sub-Technologies</td>
<td>2,239,488 Outcomes</td>
<td>3780 Scenarios</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D

UNCERTAINTY IN CLIMATE DAMAGES

A major source of uncertainty in climate change is the severity of resulting climate damages. To understand how uncertainty in climate damages affects the optimal near term responses Baker & Solak (2011, 2013) model climate damage uncertainty using mean preserving spreads. The base, no-uncertainty case is taken from (Nordhaus 2008) and assumes a 1.1% loss in GDP given a 2°C rise in mean atmospheric temperature. We consider Medium and High Risk cases, described in Table D.1. We implement uncertainty with learning: near term decisions are made under uncertainty; later decisions are made based on the revealed state of the world.

Table D.1: Damage Risk Cases. The table shows the percentage GDP loss for a 2°C rise in mean atmospheric temperature for each of the risk cases.

The Medium and High Risk cases are modeled as mean preserving spreads of the No Damage Risk case with the appropriate probabilities shown in the second row. \( \pi \) denotes the calibrated parameter from the damage equation that corresponds to each of the GDP losses.

<table>
<thead>
<tr>
<th></th>
<th>No Risk</th>
<th>Medium Risk</th>
<th>High Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP Loss (%)</td>
<td>1.1</td>
<td>0</td>
<td>3.3</td>
</tr>
<tr>
<td>Probability (%)</td>
<td>100</td>
<td>66.7</td>
<td>33.3</td>
</tr>
<tr>
<td>( \Pi )</td>
<td>0.003</td>
<td>0</td>
<td>0.009</td>
</tr>
</tbody>
</table>
APPENDIX E

EVALUATING ENERGY SCENARIO PROBABILITY

This section discusses the evaluation of the probability of an energy scenario. As a reminder, an energy scenario is a specific set of R&D targets for all the technologies. An R&D outcome \( \eta \) is a specific set of R&D performance specifications for all the sub-technology projects, the set of all R&D outcomes is given by \( \Gamma \). As each technology has several sub-technology projects, multiple R&D outcomes can result in the same energy scenario, as R&D in several sub-technology projects in the same technology can result in the same R&D specification. The probability of an R&D outcome is the product of the probability of success of each of the constituent sub-technology projects resulting in a particular R&D target, while the probability of an energy scenario is the sum of the probability of all the R&D outcomes that would result in the specific energy scenario.

\[
p_s(\zeta) = \sum_{\eta} \left( p_s(\eta_{j:(\eta \in \Gamma) \rightarrow \zeta}) \right) \quad \forall \zeta \quad (E.1)
\]

\[
p_s(\eta) = \prod_{j=1}^{J} p_s(\eta_j(x_{jk})) \quad \forall \eta \quad (E.2)
\]

Equation (E.1) evaluates the probability of a specific energy scenario, \( \zeta \). The probability of a specific energy scenario is the sum of the probability of all the R&D outcomes that map, \( f : (\eta \in \Gamma) \rightarrow \zeta \), to the energy scenario. In the event that R&D in multiple sub-technology projects are successful, only the best R&D target is assumed to diffuse into the economy.

Equation (E.2) evaluates the probability of an R&D outcome. The equation evaluates the probability of the R&D outcome \( p_s(\eta) \) as the product of the probability of success of each of the constituent sub-technology projects \( j \in J \) resulting in a
particular R&D target, $\eta_j(x_{jk})$, given that each sub-technology project is funded at level $x_{jk}$ of R&D funding.
APPENDIX F

OPTIMAL R&D PORTFOLIO

Table F.1 below shows the optimal allocation of R&D funds to the sub-technologies in the No Damage Risk case while Figures Figure F.1 and Figure F.2 show the optimal allocation of R&D funds to the sub-technologies in the Medium and High Damage Risk case respectively.
Table F.1: Optimal Portfolio for the No Damage Risk Case ($ million R&D).

*Green shade indicate the first time maximum investment is reached*

<table>
<thead>
<tr>
<th>BE</th>
<th>BF</th>
<th>CCS</th>
<th>NUC</th>
<th>SOL</th>
<th>TRN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Steam</td>
<td>Gas</td>
<td>STP1</td>
<td>STP2</td>
</tr>
<tr>
<td>300</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>400</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>500</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>600</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>700</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>800</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>1100</td>
<td>0</td>
<td>241</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1400</td>
<td>0</td>
<td>241</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1500</td>
<td>0</td>
<td>241</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1600</td>
<td>0</td>
<td>319</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1800</td>
<td>0</td>
<td>319</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1900</td>
<td>0</td>
<td>772</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2200</td>
<td>0</td>
<td>772</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2300</td>
<td>0</td>
<td>772</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2500</td>
<td>0</td>
<td>772</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2600</td>
<td>0</td>
<td>772</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2700</td>
<td>0</td>
<td>346</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3400</td>
<td>0</td>
<td>346</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4000</td>
<td>0</td>
<td>386</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4200</td>
<td>0</td>
<td>386</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4400</td>
<td>0</td>
<td>386</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4900</td>
<td>0</td>
<td>386</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6100</td>
<td>0</td>
<td>386</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7500</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7900</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10200</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11000</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12500</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14000</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18700</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29500</td>
<td>0</td>
<td>386</td>
<td>1293</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
**Figure F.1:** Optimal R&D Portfolio to Utility chart across all the R&D budget levels when there is Medium uncertainty in climate damages.

**Figure F.2:** Optimal R&D Portfolio to Utility chart across all the R&D budget levels when there is High uncertainty in climate damages.
APPENDIX G

CONSISTENCY OF THE SOLUTION

In this section, we discuss the consistency of the algorithm. Given the unusually long computational time of the implementation even with parallelization, we run the algorithm ten times for each budget level. Figure G.1 shows the variability in the portfolios returned as optimal by the genetic algorithm, where variability is the number of times the genetic algorithm returns the same optimal solution of the 10 runs.

We observe that the algorithm variability reduces as the budget level constraint increases. As the genetic algorithm searches using guided randomized search, we would expect the algorithm’s solution variability to increase as the budget constraint increases, increasing the scale of the problem. Our results are therefore contrary to intuition. This is because most of the funding levels fit within the R&D budget at higher budget levels, leading to less constraint violation and a more exhaustive search by the genetic algorithm.

![Figure G.1: Variability of the genetic algorithm solution.](image)

Figure G.1: Variability of the genetic algorithm solution.
APPENDIX H

SUB-COMPARTMENTALIZATION

Change in the Optimal Portfolio in the 5 technology problem with respect to Nuclear: Figure H.1 shows the additional benefit of having Nuclear in the portfolio. The figure shows the change in optimal portfolio when the allocation to Nuclear energy in the six technology problem is excluded from the R&D budget and the portfolio is re-optimized for the five technology problem. Again we notice a very significant, change in the composition of the optimal portfolio. From Figure H.1, we note that the presence of Nuclear in the portfolio favors the inclusion of the Batteries for Electric Transportation and Bio-Fuels energy technologies at the expense of the Solar and Bio-Electricity energy technologies. This again supports findings from Table 3 which show that Bio-Electricity and Solar are significant substitutes with Nuclear energy and a slight complement with Batteries for Electric Transportation. Therefore the presence of Nuclear leads to less funding for Solar energy.
Figure H.1: Change in optimal portfolio with respect to Nuclear.

*It shows the difference in the optimal portfolio in the six technologies problem minus the optimal portfolio in the five technologies problem.*

*Optimal Portfolio in the 5 technology problem when the optimal allocation to Bio-Electricity is excluded:* Similar to the above figures, Figure H.2 shows the difference in the sub-problem optimal portfolio relative to the Bio-Electricity energy technology.

We note that the presence of Bio-Electricity in the portfolio favors the inclusion of Carbon Capture and Storage at the expense of Bio-Fuels. This also supports findings from Table 3 which show that Bio-Electricity and Carbon Capture and Storage are significant complements, however, we note that we do not see a change in the funding allocation for Nuclear, even though Bio-Electricity is a significant substitute for it.
Figure H.2: Change in optimal portfolio with respect to Bio-Electricity.

It shows the difference in the optimal portfolio in the six technologies problem minus the optimal portfolio in the five technologies problem.

Optimal Portfolio in the 4 technology problem with CCS and Bio-Electricity excluded: Given that the regression analysis of the interactions characteristics, Table 3, show that all third-order interaction effects are insignificant, we also sub-compartmentalize the problem based on a pair of energy technologies, the CCS and Bio-Electricity energy pair. We find from Figure H.3 that the presence of R&D funding in CCS and Bio-Electricity in the funding portfolio favors the R&D development of the Bio-Fuels energy technology in lieu of Batteries for Electric Transportation. This relationship exhibited might be due to the fact that the Bio-Fuels technology, similar to CCS and Bio-Electricity, becomes more favorable as the abatement level increases relative to the other technologies.
Figure H.3: Change in optimal portfolio for some selected budget levels when CCS and Bio-Electricity is excluded.

*It shows the difference in the optimal portfolio in the six technologies problem minus the optimal portfolio in the four technologies problem.*
APPENDIX I

BIO-FUELS TECHNOLOGICAL PATHS

The Bio-fuels energy technology describes the generation of liquid fuels, including ethanol, diesel or gasoline, from biomass. Our portfolio optimization model requires that the R&D technological process be independent for each of the considered sub-technologies. We discuss the re-modification of these sub-technologies and endpoints into four consistent and independent sub-technologies, each with independent endpoints and R&D technology paths. We equally conduct an analysis of the elicited funding levels to obtain the best four non-dominated funding amounts per sub-technology.

I.1 Bio-Fuels Technological Paths and Funding Levels

A brief review of the Bio-Fuels technological paths is provided in this section. The bio-mass feedstock to bio-fuels conversion process usually consists of two stages: the first stage involves breaking down the bio-mass feedstock into simpler by-products and the second stage involves converting the broken down by-products to liquid fuel. Figure I.1 and Table I.1 give the originally assessed (Baker and Keisler 2011) technological paths and funding levels from stages 1 and 2 to the final fuel produced.

As seen in Figure I.1 there are seven possible successful R&D paths from any stage 1 process to a final liquid fuel. The paths culled from (Baker and Keisler 2011) are itemized below:

- Pyrolysis followed by Bio-Oil refining to result in either Gasoline or Diesel
- Liquefaction followed by Bio-Crude refining to result in either Gasoline or Diesel
- Hydrolysis followed by aqueous-processing to produce Diesel
- Hydrolysis followed by Fermentation to ethanol
- Gasification followed by SynGas conversion to produce Ethanol
- Gasification followed by SynGas conversion to produce Ethanol

We notice that multiple paths can lead to multiple endpoints, leading to dependence between the different sub-technologies. As an example, the Hydrolysis path can either be followed by the Aqueous Processing phase or the Fermentation phase to result in Diesel or Ethanol, respectively.

Figure I.1: Bio-Fuels technological paths from feedstock to end product.

Table I.1: Bio-Fuels Funding Trajectories.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Low Funding</th>
<th>Medium Funding</th>
<th>High Funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Pyrolysis</td>
<td>$2.5M 10 years</td>
<td></td>
<td>$67.5M 10 years</td>
</tr>
<tr>
<td>2 Liquefaction</td>
<td>$2.5M 10 years</td>
<td></td>
<td>$67.5M 10 years</td>
</tr>
<tr>
<td>3 Hydrolysis</td>
<td>$0 beyond current</td>
<td>$20M 7 years</td>
<td>$270M 10 years</td>
</tr>
<tr>
<td>4 Gasification</td>
<td>$5M 10 years</td>
<td>$175M 10 yrs</td>
<td>$175M 10 years + High pyrolysis and Liquefaction</td>
</tr>
<tr>
<td>5 Refine bio-oil</td>
<td>$10M 10 years</td>
<td></td>
<td>$100M 10 years</td>
</tr>
<tr>
<td>6 Refine bio-crude</td>
<td>$10M 10 years</td>
<td></td>
<td>$100M 10 years</td>
</tr>
<tr>
<td>7 A-P Processing</td>
<td>$3M 10 years</td>
<td></td>
<td>$40M 10 years</td>
</tr>
<tr>
<td>8 Fermentation</td>
<td>High for A-P only</td>
<td></td>
<td>High A-P + $10M 10 years</td>
</tr>
<tr>
<td>9 Syngas to Diesel</td>
<td>$2M 10 years</td>
<td></td>
<td>$4M 10 years</td>
</tr>
<tr>
<td>10 Syngas to Ethanol</td>
<td>$2M 10 years</td>
<td></td>
<td>$4M 10 years</td>
</tr>
</tbody>
</table>
I.1.1  Funding Levels

A summary of the aggregated probabilities from the expert elicitations from (Baker and Keisler 2011) for the R&D funding levels is given in Table I.2. The Hydrolysis, Gasification and Fermentation processes are elicited at three funding levels while all the other paths are elicited at only the low and high funding levels, in addition to the No funding level. As each of the technological paths consists of at least two processes, a problem exists in identify the funding levels that are best. We discuss the identification of these funding levels in section I.3.

Table I.2: Aggregated Probabilities for all Bio-Fuels Paths.

<table>
<thead>
<tr>
<th>Path</th>
<th>Low Funding</th>
<th>Med Funding</th>
<th>High Funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyrolysis</td>
<td>63%</td>
<td></td>
<td>91%</td>
</tr>
<tr>
<td>Liquefaction</td>
<td>36%</td>
<td></td>
<td>61%</td>
</tr>
<tr>
<td>Hydrolysis</td>
<td>27%</td>
<td>30%</td>
<td>66%</td>
</tr>
<tr>
<td>Gasification</td>
<td>12%</td>
<td>66%</td>
<td>82%</td>
</tr>
<tr>
<td>Refine bio-oil</td>
<td>40%</td>
<td></td>
<td>83%</td>
</tr>
<tr>
<td>Refine bio-crude</td>
<td>46%</td>
<td></td>
<td>93%</td>
</tr>
<tr>
<td>A-P Processing</td>
<td>31%</td>
<td></td>
<td>58%</td>
</tr>
<tr>
<td>Fermentation</td>
<td>64%</td>
<td>67%</td>
<td>73%</td>
</tr>
<tr>
<td>Syngas to Diesel</td>
<td>50%</td>
<td></td>
<td>57%</td>
</tr>
<tr>
<td>Syngas to Ethanol</td>
<td>38%</td>
<td></td>
<td>47%</td>
</tr>
</tbody>
</table>

I.2  Making technological paths independent

The steps taken to re-modify the Bio-fuels technological paths to make them independent are discussed in this section and are shown in Figure I.2

Selective Thermal Processing (Pyrolysis 1 and Liquefaction 2) sub- technologies/paths: The selective thermal processing paths are independent of each other. Pyrolysis can be followed by bio-oil refining and Liquefaction is followed by bio-
crude refining. Therefore, both sub-technologies are not modified, and are defined as independent sub-technologies leading to either a low or high success endpoint (Diesel or Gasoline). As the sub-technological paths consist of two sub-technology projects, we need to combine the assessed success probabilities.

**Hydrolysis Sub-technology**: There are three possible paths from hydrolysis to a liquid fuel as originally defined, Figure I.1. One is hydrolysis followed by aqueous phase processing to obtain diesel, another is hydrolysis followed by aqueous phase processing to obtain ethanol as the end-product, and the third is hydrolysis followed by fermentation to obtain ethanol. As these are not independent, we collapse the paths into one technological path to fit within our portfolio optimization structure. We define the Hydrolysis path as hydrolysis followed by both aqueous phase processing and fermentation. The path can result in either of two end-points: diesel or ethanol. We omit the 77% efficiency Aqueous Processing stage as most of the experts thought that this was unrealistic (Baker and Keisler 2011). The combination of the funding levels and elicited probabilities of the sub-technologies are defined in section I.3.

**Gasification Sub-technology**: The gasification path can result in either of two endpoints; ethanol or diesel. This technological path is not modified and is assessed as leading to either a Low or High endpoint.
Choosing appropriate funding levels

As previously discussed, each of the technological paths consists of at least two technological processes (shown in Figure I.1), with each of the technological processes elicited at different levels of funding (shown in Table I.2). As an example, the hydrolysis sub-technology will have 48 possible R&D funding levels/success probabilities as the hydrolysis process was elicited at the four funding levels, aqueous phase processing at three funding levels and fermentation at four funding levels. Our portfolio model however restricts us to three funding levels (High, Medium and Low) in addition to the reference No R&D funding level. We therefore discuss the identification of these Pareto optimal funding level success probability combinations in this section.
We carry out this pre-optimization of the funding levels using the expected utility of each of the technological endpoints\(^7\) obtained from the DICE model. The alternate possible funding levels are compared by their probability weighted utility to R&D funding ratio across all the possible endpoints that can result. We ensure the constituent processes in each technological path are all funded at the appropriate levels to ensure the subsequent success of the subsequent processes. We discuss the specifics in the following sections.

### I.3.1 Selective Thermal Processing 1 & 2 sub-technologies

The two technological paths that make up the STP1 sub-technology are independent by definition. Figure I.3 and Figure I.4 show the R&D funding levels against the probability weighted utility of R&D success for all the endpoints that can result. As an example, the Low Low point is the sum of the net present value of the Low Pyrolysis funding level and the Low Bio-Oil refining funding level. The probability of the Low Low point is the product of the probability of success of the Pyrolysis and Bio-Oil refining when funded at the low funding level. The probability weighted funding is then the product of the probability of the Low Low funding point and the utility of the Gasoline endpoint. Three of the best funding level combinations are then selected as shown in Figure I.3. The same approach is used in the Liquefaction and Bio-Crude refining path.

\(^7\) Where an endpoint is a group of cost, efficiency and technical improvements target for a particular sub-technology
Figure I.3: STP1 Best Budgets to Weighted Expected Utility.

Figure I.4: STP2 Best Budgets to Weighted Expected Utility.

I.3.2 Hydrolysis sub-technology:
As discussed above in Figure I.1, this sub-technology consists of the Hydrolysis process followed by either Aqueous phase processing or Fermentation. Due to success restrictions from the Fermentation (Baker and Keisler 2011) and Table I.1 we assume the Hydrolysis is a serial path, starting with the Hydrolysis process followed by the Aqueous Phase Processing and then the Fermentation process. Thus there are four different events that can happen in the case that all the 3 processes are funded, these are:

1. The three processes are successful
2. All the processes are successful except fermentation
3. All the processes are successful except aqueous phase processing

There are also two endpoints that can result from this endpoint in addition to the reference no development endpoint (Diesel and Ethanol). Note that we have ignored the possibility of aqueous phase processing leading to the ethanol endpoint. Due to the nature of the dependent technological paths as defined in (Baker and Keisler 2011) and Figure I.1, low and high funding in fermentation is possible only if A-P processing has been invested in at the high funding level. The result of the weighted expected utility analysis for each of the funding levels is shown in Figure I.5.

Three non-dominated funding levels based on their expected weighted utility to R&D funding ratio are chosen as seen in Figure I.5.
Figure I.5: Hydrolysis sub-technology Best Budgets to Weighted Expected Utility.

### I.3.3 Gasification

The gasification sub-technology consists of the first stage gasification path followed by either the ‘synthetic to diesel’ conversion or the ‘synthetic to ethanol’ conversion. The four possible events that can result if all the paths are funded are:

1. All the three paths are successful
2. All are successful but the synthetic to diesel conversion fails
3. All are successful but the synthetic to ethanol conversion
4. Gasification fails and or all paths fail.
There are two possible endpoints in addition to the failure or reference endpoint.

Figure I.6 shows the selected three non-dominated funding levels based on the weighted expected utility to the R&D funding level.

![Figure I.6: Gasification sub-technology Best Budgets to Weighted Expected Utility.](image)

**I.4 Final bio-fuels redefined sub-technologies elicitations and funding levels**

The final redefined success probabilities and funded levels after the requirements for technological path independence and compactness are enforced are shown in Table I.3.
Table I.3: Bio-Fuels funding levels, success probabilities and cost endpoints

<table>
<thead>
<tr>
<th>Technology</th>
<th>Sub Technology</th>
<th>Funding ($M)</th>
<th>Success Probability</th>
<th>Cost Endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-Fuels</td>
<td>STP1</td>
<td>1293</td>
<td>76%</td>
<td>STP1 Endpoint</td>
</tr>
<tr>
<td></td>
<td></td>
<td>791</td>
<td>52%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>97</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STP2</td>
<td>1293</td>
<td>57%</td>
<td>STP2 Endpoint</td>
</tr>
<tr>
<td></td>
<td></td>
<td>791</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>97</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HYDROLYSIS</td>
<td>2471</td>
<td>48%</td>
<td>High-Ethanol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10%</td>
<td>Low-Diesel</td>
</tr>
<tr>
<td></td>
<td></td>
<td>77</td>
<td>18%</td>
<td>High-Ethanol</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>8%</td>
<td>Low-Diesel</td>
</tr>
<tr>
<td></td>
<td>GASIFICATION</td>
<td>1413</td>
<td>38%</td>
<td>High-Diesel</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13%</td>
<td>Low-Ethanol</td>
</tr>
<tr>
<td></td>
<td></td>
<td>85</td>
<td>7%</td>
<td>High-Diesel</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2%</td>
<td>Low-Ethanol</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
<td>6%</td>
<td>High-Diesel</td>
</tr>
</tbody>
</table>
APPENDIX J

GLOBAL SENSITIVITY ANALYSIS METHODS

The following figures (Figure J.1, Figure J.2 and Figure J.3) show the cumulative distribution function of the CDF based GSA Kolmogorov-Smirnov test for the Labor Productivity Growth, Income Elasticity of Demand and AEII respectively, relative to the radiative forcing output.

Figures (Figure J.4, Figure J.5, Figure J.6) provide the conditional density function for the Borgonovo Delta Density based Global Sensitivity Analysis for the three inputs.

Figure J.1: Cumulative distribution function of the labor productivity growth relative to radiative forcing using the distance to CDF based Kolmogorov Smirnov test.
Figure J.2: Cumulative distribution function of the income elasticity of demand relative to radiative forcing using the distance to CDF based Kolmogorov Smirnov test.

Figure J.3: Cumulative distribution function of the AEII relative to radiative forcing using the distance to CDF based Kolmogorov Smirnov test.
Figure J.4: Density function of the conditional labor productivity growth relative to radiative forcing using the density based Borgonovo Importance measure.

Figure J.5: Density function of the conditional income elasticity of demand relative to radiative forcing using the density based Borgonovo Importance measure.
Figure J.6: Density function of the conditional AEII relative to radiative forcing using the density based Borgonovo Importance measure.
BIBLIOGRAPHY


