



# Essays on Energy Technology Innovation Policy

## Citation

Chan, Gabriel Angelo Sherak. 2015. Essays on Energy Technology Innovation Policy. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467190>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Essays on Energy Technology Innovation Policy

A dissertation presented

by

Gabriel Angelo Sherak Chan

to

The Department of Public Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Public Policy

Harvard University  
Cambridge, Massachusetts

April 2015

© 2015 Gabriel Angelo Sherak Chan

All rights reserved.

*Dissertation Advisors:*  
**Professor William Clark**  
**Professor Laura Díaz Anadón**

*Author:*  
**Gabriel Angelo Sherak Chan**

# **Essays on Energy Technology Innovation Policy**

## **Abstract**

Motivated by global climate change, enhancing innovation systems for energy technologies is seen as one of the largest public policy challenges of the near future. The role of policy in enhancing energy innovation systems takes several forms: public provision of research and develop funding, facilitating the private sector's capability to develop new technologies, and creating incentives for private actors to adopt innovative and appropriate technologies. This dissertation explores research questions that span this range of policies to develop insights in how energy technology innovation policy can be reformed in the face of climate change.

The first chapter of this dissertation explores how decision making to allocate public research and development funding could be improved through the integration of expert technology forecasts. I present a framework to evaluate and optimize the U.S. Department of Energy's research and development portfolio of applied energy projects, accounting for spillovers from technical complimentary and competition for the same market share. This project integrates one of the largest and most comprehensive sets of expert elicitations on energy technologies (Anadón et al., 2014b) in a benefit evaluation framework. This work entailed developing a new method for probability distribution sampling that accommodates the information that can be provided by expert elicitations. The results of this project show that public research and development in energy storage and solar photovoltaic technologies

has the greatest marginal returns to economic surplus, but the methodology developed in this chapter is broadly applicable to other public and private R&D-sponsoring organizations.

The second chapter of this dissertation explores how policies to transfer technologies from federally funded research laboratories to commercialization partners, largely private firms, create knowledge spillovers that lead to further innovation. In this chapter, I study the U.S. Department of Energy's National Laboratories, and provide the first quantitative evidence that technology transfer agreements at the Labs lead to greatly increased rates of innovation spillovers. This chapter also makes a key methodological contribution by introducing a technique to utilize automated text analysis in an empirical matching design that is broadly applicable to other types of social science studies. This work has important implications for how policies should be designed to maximize the social benefits of the \$125 billion in annual federal funding allocated to research and development and the extent to which private firms can benefit from technology partnerships with the government.

The final chapter of this dissertation explores the effectiveness of international policy to facilitate the deployment of low-emitting energy technologies in developing countries. Together with Joern Huenteler, I examine wind energy deployment in China supported through international climate finance flows under the Kyoto Protocol's Clean Development Mechanism. Utilizing a project-level financial model of wind energy projects parameterized with high-resolution observations of Chinese wind speeds, we find that the environmental benefits of projects financed under the Clean Development Mechanism are substantially lower than reported, as many Chinese wind projects would have been built without the Mechanism's support, and thus do not represent additional clean energy generation.

Together, the essays in this dissertation suggest several limitations of energy technology innovation policy and areas for reform. Public funds for energy research and

development could be made more effective if decision making approaches were better grounded in available technical expertise and developed in framework that captures the important interactions of technologies in a research and development portfolio. The first chapter of this dissertation suggests a politically feasible path towards this type of reform.

Policies to “unlock” publicly sponsored inventions from the organizations that develop them have broad impact on private sector innovation. These policies multiply the effect of public research and development funds, but should be strengthened to more rapidly advance the scientific frontier. The second chapter of this dissertation provides some of the first quantitative evidence to support reform in this area.

Finally, international policies to facilitate the deployment of climate-friendly technologies in developing countries face serious implementation challenges. The current paradigm of utilizing carbon markets to fund individual projects that would not have otherwise occurred has failed to encourage energy technology deployment in one of the sectors with the greatest experience with such policies. The third chapter of this dissertation suggests that this failure has been largely due to poorly designed procedural rules, but options for reform are available.

Mitigation of global climate change will require broad policy response across the full range of scales, sectors, and policy spheres. Undoubtedly, climate mitigation will result in widespread transformation of energy systems. This dissertation focuses on the role of innovation policy in accelerating the transformation of these systems. The range of policies studied in this dissertation can make climate change mitigation more politically feasible and more cost effective by expanding the set of technological choices available to public and private actors faced with incentives and requirements to lower their greenhouse gas emissions to collectively safe levels.

# Contents

Abstract .....	iii
Acknowledgments.....	xii
1 Improving Decision Making for Public R&D Investment in Energy: Utilizing Expert Elicitation in Parametric Models.....	1
1.1 Introduction.....	2
1.1.1 Current Process at the U.S. Department of Energy .....	3
1.1.2 Design Principles for an R&D Decision-Support Tool .....	5
1.1.3 Proposed Methods in the Literature .....	10
1.1.4 The Challenges of R&D Benefit Estimation .....	11
1.2 Methods .....	13
1.2.1 Expert Elicitation .....	14
1.2.2 Formalizing Elicitation Results.....	18
1.2.3 Sampling Strategy .....	21
1.2.4 Modeling with MARKAL .....	27
1.2.5 Optimization .....	29
1.3 Results .....	34
1.3.1 System Benefits .....	34
1.3.2 Technology Diffusion .....	35
1.3.3 Uncertainty Analysis .....	37
1.3.4 Optimized R&D Portfolios .....	39
1.4 Conclusions and Discussion .....	43
1.5 Acknowledgments .....	46
2 The Commercialization of Publicly Funded Science: How Licensing Federal Laboratory Inventions Affects Knowledge Spillovers .....	47
2.1 Introduction.....	48

2.2	The U.S. National Labs and Technology Transfer .....	54
2.2.1	Lessons from University Technology Transfer .....	58
2.2.2	National Lab Patent Licenses .....	59
2.3	Data .....	62
2.4	Estimating a Citation-Based Model of Knowledge Diffusion.....	65
2.4.1	Matching .....	66
2.4.2	Regression Framework .....	85
2.5	Results .....	89
2.5.1	Diffusion after Licensing .....	89
2.5.2	Exclusive versus Non-Exclusive Licenses .....	98
2.5.3	Breadth versus Concentration of Spillovers .....	100
2.5.4	Accounting for Strategic Patenting and Signaling .....	101
2.6	Discussion and Conclusion.....	103
2.7	Acknowledgments .....	106
3	Financing Wind Energy Deployment in China through the Clean Development Mechanism .....	108
3.1	Introduction.....	109
3.1.1	Offsets in Theory.....	110
3.1.2	Offsets in Practice: The Clean Development Mechanism.....	117
3.1.3	The Chinese Wind Sector .....	120
3.2	The Role of the CDM in Financing Wind Energy in China.....	124
3.2.1	Introduction .....	124
3.2.2	Estimating the Share of CDM Projects in China’s Wind Sector.....	126
3.2.3	Results.....	131
3.2.4	Discussion .....	140
3.2.5	Conclusions and Policy Implications.....	143
3.3	The Additionality of Clean Development Mechanism Projects in the Chinese Wind Sector .....	145
3.3.1	Introduction .....	147
3.3.2	Data.....	151
3.3.3	Methods and Results .....	159
3.3.4	Discussion .....	166
3.3.5	Conclusions and Policy Implications.....	167
3.4	Acknowledgments .....	170



A	Appendix to Chapter 1 .....	187
	A.1 Expert Elicitation Protocol.....	187
	A.1.1 Technology Selection.....	187
	A.1.2 Expert Selection.....	188
	A.1.3 Expert Elicitation Protocol .....	188
	A.1.4 Qualitative Reviews of Elicitation Results.....	194
	A.2 Correlation of Future Technology Cost Improvements .....	194
B	Appendix to Chapter 2 .....	196
	B.1 The Latent Dirichlet Allocation Model .....	196
C	Appendix to Chapter 3 .....	200
	C.1 Internal Rate of Return for a Wind Farm .....	200
	C.2 Chinese Policies Affecting Wind Energy Development.....	203

## List of Tables

Table 1.1: Summary of Four Design Principles for an R&D Decision Making Process.....	10
Table 2.1: Summary Statistics of R&D, Patenting, and Licensing by Lab Operator .....	56
Table 2.2: Descriptive Statistics for All Patents and Patents that are Ever Licensed.....	64
Table 2.3: Conditional Means of Annual Citations Approximating Diff-in-Diff.....	77
Table 2.4: Balance Check .....	78
Table 2.5: Sensitivity Analysis with Different Matching Approaches .....	81
Table 2.6: Baseline Regression Estimates .....	92
Table 2.7: Robustness to Functional Form .....	94
Table 2.8: Robustness to Matching Design.....	96
Table 2.9: Exclusive versus Non-Exclusive Licenses.....	99
Table 2.10: Concentration of Diffusion .....	101
Table 2.11: Accounting for Strategic Patenting.....	103
Table 3.1: Estimates of the Share of Chinese Wind Projects that Received CDM .....	127
Table 3.2: Technology Transfer Classifications .....	130
Table 3.3: Cross Validation Functional Forms .....	153
Table 3.4: Cross Validation Model Evaluation .....	155
Table 3.5: Annual Difference in IRR for CDM and Non-CDM Projects .....	164
Table A.1: Estimated Cross-Technology Spearman Correlations .....	195
Table C.1: IRR Model Inputs and Assumptions .....	203
Table C.2: Key Policies and Plans Affecting Onshore Wind Power in China 1994-2011 ...	204

## List of Figures

Figure 1.1: Comparison of Expert-Recommended R&D Levels.....	17
Figure 1.2: Distribution-Fitting to Expert-Elicited Percentiles .....	23
Figure 1.3: An Example Sample from the Joint Distribution of Five Technology Costs.....	25
Figure 1.4: Graphical Depiction of Dependent Sampling.....	27
Figure 1.5: Summary of Methodology from Elicitations to an Optimal Portfolio .....	33
Figure 1.6: Distribution of Benefits under Two R&D Scenarios .....	38
Figure 1.7: Optimal R&D Portfolios under a No Climate Policy Scenario.....	41
Figure 1.8: Optimal Allocation and Marginal Returns by Technology Area .....	42
Figure 2.1: Time Series of Federal Funding for R&D.....	49
Figure 2.2: Location of the 17 U.S. National Labs under DOE.....	55
Figure 2.3: Patenting and Licensing of Federally Funded R&D by Sponsoring Agency.....	57
Figure 2.4: Distribution in the Lag between Patent Filing and Licensing .....	62
Figure 2.5: Example Classification of a Patent using the LDA Model.....	72
Figure 2.6: Estimated Hazard Ratios for Each Topic.....	75
Figure 2.7: Annual Difference-in-Difference Estimates .....	93
Figure 2.8: Sensitivity of Coefficient Estimates under Different Matching Procedures .....	97
Figure 2.9: Sensitivity of Coefficient Estimates under Different Matching Procedures .....	98
Figure 3.1: The Supply and Demand for Projects that Qualify for an Offset Scheme.....	114
Figure 3.2: Transaction Costs and Environmental Integrity in Offset Schemes.....	117
Figure 3.3: Geographic Distribution of Investment in CDM Projects.....	119
Figure 3.4: Distribution of Chinese CDM Projects by Project Class.....	120
Figure 3.5: Cumulative Installed Capacity of Wind Power by Country.....	122
Figure 3.6: Spatial Distribution of Chinese Wind Farms by Province.....	123
Figure 3.7: Installed Wind Capacity in China under the CDM .....	133
Figure 3.8: Capital Investment under the CDM in the Chinese Wind Sector .....	134

Figure 3.9: CER Generation and Revenue by Chinese Wind CDM Projects .....	136
Figure 3.10: Differences in Technology Choice between CDM and non-CDM Projects.....	138
Figure 3.11: Average Distance of New Wind Projects to Closest Existing Projects .....	139
Figure 3.12: Cumulative Installed Wind Capacity in China from 1998 – 2014.....	146
Figure 3.13: Self-Reported Internal Rates of Return for Chinese Wind CDM Projects .....	150
Figure 3.14: Capacity Factor Estimates and Observations.....	156
Figure 3.15: Comparison of Capacity Factor Estimates and Observations .....	157
Figure 3.16: Modeled Investment Cost and Tariff versus Observations .....	159
Figure 3.17: Distribution of Calculated IRR for CDM and Non-CDM Projects .....	165
Figure A.1: Schematic of the Protocol for Expert Elicitations.....	189

## Acknowledgments

Writing my dissertation was a blast, but only because of the amazing guidance and support I've had.

First, I would like to thank my inspiring advisors, Bill Clark, Laura Díaz Anadón, and Joe Aldy. Their guidance has pushed my thinking beyond what I could have imagined when I started grad school, and my work would be radically different without their support. I owe a particular debt to Laura who took me under her wings at the very beginning of my Ph.D. and has shared seemingly endless amounts of her time and attention with me.

I also owe a great deal to one of my earliest mentors, Bob Marlay. The two summers I spent working for Bob at the Department of Energy inspired what has now become a career-long passion for energy innovation and climate change. Thank you for planting so many seeds in my brain back in 2007 and 2009. Thank you also to my undergraduate mentors at M.I.T. who helped me taste the first flavors of research: Ron Prinn, Ken Oye, John Reilly, and Sergey Paltsev.

I am deeply grateful to a wide range of informal advisors and cheerleaders at Harvard who have challenged me and supported me in my intellectual growth: Venky Narayanamurti, F.M. Scherer, Richard Freeman, Rob Stavins (who took me around the world for the past four years to work on the IPCC), Bill Hogan, Henry Lee, Nancy Dickson, Karin Vander Schaff, Jason Chapman, and Rob Stowe.

My colleagues and friends at Harvard and beyond have been tremendous in getting me through the Ph.D. process smarter and saner. Thank you to Rich Sweeney, Trisha Shrum, Tara Grillos, Alicia Harley, Dan Honig, Ariel Dora Stern, Liz Walker, Steffen Jenner, Joern

Huenteler, Nidhi Santen, Audrey Lee, Amitai Bin-Nun, Sarah Jane Maxted, Neil Chao, and Zach Shapiro.

I couldn't imagine the last six years without the people in the research groups that I've been fortunate enough to participate in. The Energy Technology Innovation Policy Group was my dream research home when I was applying to grad school and it lived up to everything I could have imagined. I've met amazing colleagues there to whom I owe a great debt. The Belfer Center for Science and International Affairs, the Harvard Environmental Economics Group, and the Sustainability Science Program have given me tremendous resources to do the research I wanted to do while expanding my horizons. I am grateful for their financial support.

Finally, thank you to my family. My parents, Kathy Sherak and Chaen Chan, taught me with their actions that the best way to be happy in life is to find something to be passionate about and then never look back. Very special thanks to Stephanie Au for her limitless support and unconditional, tougher-than-nails love.

## **Chapter 1**

# **Improving Decision Making for Public R&D Investment in Energy: Utilizing Expert Elicitation in Parametric Models**

Effective decision making to allocate public funds for energy technology research, development, and demonstration (R&D) requires considering alternative investment opportunities that can have large but highly uncertain returns and a multitude of positive or negative interactions. This chapter proposes and implements a method to support R&D decisions that propagates uncertainty through an economic model to estimate the holistic benefits of an R&D portfolio, accounting for innovation spillovers and technology substitution and complementarity. The proposed method improves on the existing literature by: (a) using estimates of the impact of R&D investments from the most comprehensive set of expert elicitations on this topic to date; (b) using a detailed

energy-economic model to estimate evaluation metrics relevant to an energy R&D portfolio: e.g., system benefits, technology diffusion, and uncertainty around outcomes; and (c) using a novel sampling and optimization strategy to calculate optimal R&D portfolios. This design is used to estimate an optimal energy R&D portfolio that maximizes the net economic benefits under an R&D budget constraint. Results applied to the United States indicate that: (1) the median projection of the expert-recommended portfolio in 2030, relative to the BAU portfolio, reduces carbon dioxide emissions by 46 million tonnes, increases economic surplus by \$29 billion, and increases renewable energy generation by 39 TWh; (2) uncertainty around the estimates of R&D benefits is large and overall uncertainty increases with greater investment levels; (3) a 10-fold expansion from 2012 levels in the R&D budget for utility-scale energy storage, bioenergy, advanced vehicles, fossil energy, nuclear energy, and solar photovoltaic technologies is justified by returns to economic surplus; (4) the greatest returns to public R&D investment are in energy storage and solar photovoltaics; and (6) the current allocation of energy R&D funds is very different from optimal portfolios. Taken together, these results show that reform of public energy R&D decision making processes can have an important impact on the cost-effectiveness of R&D and carbon mitigation efforts.

## **1.1 Introduction**

The well-known environmental (IPCC, 2014), economic (IEA, 2011; WHO and UNDP, 2009), and security (Cherp et al., 2012) challenges in the energy sector have justified a wide array of energy policies, such as pollution regulations, targeted subsidies, and technology standards. In order to be dynamically cost-effective, these policies require complementary technology policies (Jaffe et al., 2005). One of the most important forms of technology policy is government funding for energy R&D to support innovation



projects that would not otherwise attract sufficient private investment. Toward this end, the U.S. Department of Energy (DOE) has allocated approximately \$5 billion per year since 2009 towards energy R&D (Gallagher and Anadón, 2013). The method presented in this chapter offers an approach to supporting decisions to allocate public R&D funds across technology areas that complements existing frameworks such as comprehensive reviews of energy R&D activities (DOE, 2012), retrospective analysis (NRC, 2001), individual technology roadmaps (Geum and Park, 2013), and evaluations driven by the broader political landscape.

The method proposed and implemented here has several desirable characteristics that differentiates it from existing approaches in the literature: (1) it quantifies the anticipated improvement in technology in a way that fully accounts for the inherent uncertainty in the returns to R&D; (2) it systematically captures the interactions between R&D investments that can occur through technology spillovers or through market interactions of technology substitutes and complements, an important limitation of existing approaches (NRC, 2007a); (3) it is flexible to changing assumptions that are fundamentally subjective, such as belief distributions of the elasticity of future technology costs to current R&D investments; and (4) it is transparent, and therefore feasible to implement in public organizations that require transparency to build procedural legitimacy.

### **1.1.1 Current Process at the U.S. Department of Energy**

DOE is the single largest energy R&D funding entity in the United States and the largest public energy R&D funding entity of all member countries in the International Energy Agency (which are largely the industrialized country members of the OECD and the countries with the most reliable data for this metric) (IEA, 2013). Nevertheless,

several expert panels have called for greater U.S. government spending in energy R&D (American Energy Innovation Council, 2010; APS, 2008; NCEP, 2004; PCAST, 2010, 1997). Although each of these studies calls for additional resources to be devoted to energy R&D, none of them offer rigorous quantitative estimates of the expected benefits of their recommendations.

In current practice, despite the demonstrated interest of outside experts to affect DOE decision-making, DOE's decision making processes and tools do not systematically consider the benefits, and the uncertainty inherent in the benefits, of individual R&D programs or in aggregate (NRC, 2007b). Current U.S. law requires that DOE submit annual estimates of its program's benefits through the Government Performance and Results Act (GPRA). However, this program does not require DOE to consider uncertainty in its evaluation of R&D programs (NRC, 2007b), nor the interactions between the different technologies that the DOE's programs support . For example, many of the technologies that DOE R&D programs support may induce technological spillovers among them or compete in markets as either complements (e.g., utility scale energy storage and renewable technologies like wind and solar power) or substitutes (e.g., more efficient internal combustion engine vehicles and electric vehicles). Capturing these market interactions is essential for estimating the benefits of improvements in technologies in a way that appropriately accounts for available "next-best" and "enabling" technologies (NRC, 2007a). Therefore, the benefits calculated under GPRA requirements are likely to be biased. Further, to improve the credibility and political buy-in to its decision-making process, DOE also faces the challenge of integrating a wide array of non-transparent technical assumptions and models promoted by various stakeholders (NRC, 2007a; Silverman, 1981).

### 1.1.2 Design Principles for an R&D Decision-Support Tool

In this section four design principles for a decision-support tool that can feasibly and effectively improve public R&D portfolio design are proposed. While these principles are broadly applicable to R&D decision making in many contexts, they were developed based on a consideration of U.S. public energy R&D decision making and through iterative discussions with analysts involved in the DOE's policymaking environment. These principles are used to evaluate how decisions are made in public energy R&D in the United States, while exploring some of the explanations for why policymakers in many public R&D funding agencies do not generally follow these design principles. The first two principles deal with analytical requirements and the second two principles concern institutional feasibility of implementing a decision support tool. Following a discussion of these principles, the remainder of the chapter presents a method that follows the four proposed design principles and then examines the results of implementing it in the case of public U.S. energy R&D.

<p><b>Principle 1: Quantifiable.</b> <i>Technological improvement benefits of a decision must be prospectively quantified and account for uncertainty</i></p>
---

This chapter differentiates “technological improvement benefits” of R&D, the change in cost and performance of technologies as a result of R&D, from “social benefits” of R&D, the changes in systemic policy objectives, such as aggregate economic surplus. Under Principle 1, technology improvement benefits are addressed; social benefits are addressed under Principle 2. The technology improvement benefits of R&D investments needed to support decision making should follow four sub-principles: (1) Relating technological improvement benefits to R&D investments requires specifying technological improvement benefits conditional on multiple levels of R&D investment;

(2) The marginal rate of technological improvements at a given level of R&D depends not only on the level of R&D in that technology area but also on the level of R&D in related technology areas. In other words, R&D-induced improvements in different technologies can be correlated due to inter-technology spillovers. Therefore, to avoid biased estimates of aggregate benefits, benefits of individual technology improvements must be jointly specified with an explicit dependence structure; (3) The returns to R&D are often only realized over long time horizons (several decades); therefore, to fully account for the benefits of R&D investments, benefit estimates must also consider both short-run and long-run technology improvement benefits; and (4) The returns to R&D should account for uncertainty. In addition to these four sub-principles, technological improvement benefit estimates should also be developed in a framework of common assumptions. For example, if estimates of the technology improvement benefits in one technology area are made for one timeframe, estimates for the benefits of other technologies should be made with the same timeframe.

In U.S. energy R&D decision making, while benefits of R&D programs are estimated conditional on R&D levels—and are occasionally considered over different time horizons—current practice neither explicitly considers the dependence in improvements between technologies nor uses common assumptions to make the estimates across technology areas. Further, current practice does not consider the uncertainty in the benefits of R&D (NRC, 2007b). Despite robust evidence that uncertainty is a defining characteristic of R&D investment, it is neither legislative requirement (e.g. through GPRA) nor standard practice (e.g. in DOE budget justification documents) to quantify or otherwise assess the uncertainty in the benefits of R&D programs.

**Principle 2: Comprehensive.** *Social benefits of R&D investments must be evaluated in a common framework*

The second principle is based on the prerequisite that sound consideration of R&D tradeoffs requires a comprehensive framework for analyzing the social benefits of various R&D investments jointly and with common metrics. Therefore, while initial technology improvement benefit estimates need not necessarily be expressed in common units, a necessary condition to evaluate tradeoffs between R&D investments is that the ultimate metric for comparison—the social benefits—should be expressed in common units.

To make well-informed R&D allocation decisions, a decision-support tool must consider how the benefits of improvements in individual technologies interact (as substitutes or complements). These interactions may be positive, as is the case with the complementary role utility-scale energy storage can play in smoothing out the intermittent supply of electricity from renewable technologies like wind and solar power, or they may be negative, as is the case with substitute technologies like efficient internal combustion engines and electric vehicles that compete for the same market share. Therefore, the benefits to society need to be estimated using a single framework that allows the aggregate benefits of a suite of R&D investments to be estimated.

In terms of the second principle, current practice in U.S. energy R&D has significant room for improvement. DOE justifications to support funding requests are typically constructed project-by-project, program-by-program, or office-by-office, with little effort to standardize assumptions or reporting metrics. As a result, recent observers have characterized DOE decision making as “being badly ‘stovepiped,’ meaning that the various offices and programs poorly communicate with one another” (Cho, 2013), and

needing a strengthened “integrated policy assessment capability” for the “analysis capabilities housed in each major program area.” (Moniz, 2013) To improve the credibility and political buy-in to its decision-making process, DOE also faces the challenge of integrating the wide array of technical assumptions and models it has access to (NRC, 2007b; Silverman, 1981). DOE’s budget justification documents contain very little information to help policymakers assess the relative merits of R&D programs, leaving decisions to be informed instead based on disparate one-dimensional estimates of benefits that do not take into account the outcomes of simultaneously-occurring decisions.

**Principle 3: Adaptive.** *Benefit analysis should be flexible to changing assumptions*

The third principle results from the need for a decision support tool to be relevant as technological characteristics improve over time and exogenous policy decisions are made (such as setting aggregate R&D budget levels or enacting policies that complement R&D). An R&D support tool should be made flexible to changing assumptions, allowing decision makers to update assumptions with the latest available information without reinventing the framework for analysis.

Flexibility to changing assumptions also has the added advantage of allowing decision makers to conduct sensitivity analysis, directly testing the effect of different assumptions. In the context of many R&D policy-making organizations, managers of individual technology programs may hold different subjective beliefs about the benefits of an R&D program. The ability to adjust to different sets of assumptions can help focus internal deliberations on specific quantitative assessments, or the choice of model and metrics to estimate social benefits, rather than abstract debates about biases of individual managers. Experience has shown that by agreeing on a flexible methodology

prior to the introduction of technical assumptions, parties can build credibility (Parson, 1998).

<p><b>Principle 4: Transparent.</b> <i>Transparency in developing assumptions and analytical methods should be feasible</i></p>
---

The fourth principle is motivated by the institutional feasibility constraint that most R&D decision making organizations face to build procedural legitimacy both internal and external to the organization. One of the most important factors in developing procedural legitimacy in this context involves managing the transparency of inputs and methods used to support decision making. Within an organization, transparency can help build credibility of estimates from managers competing for the same pool of funds who might otherwise doubt the reliability of estimates from others. Transparency in how assumptions are developed and used can also help build external (public) credibility, and therefore, political support. However, public transparency can preclude the incorporation of proprietary information, which can lower the quality of technical estimates. Therefore, a process to support R&D decision making must be feasibly transparent but not necessarily actually both publically and privately transparent.

Current U.S. energy R&D decision making practice does not make its assumptions and process transparent, making it infeasible to determine if current benefit estimates are flexible to changing assumptions or amenable to sensitivity analysis. In current practice, the technology assumptions used to estimate the social benefits of individual technology programs often come from anonymous scientists or other experts within the DOE program offices, raising questions about the independence of these benefit estimates. So long as program managers benefit from additional funding, they may suffer from motivational bias and are incentivized to overestimate the effectiveness of

their programs to increase the funding they receive. This had led to an erosion of trust among technical experts internal to R&D programs who hold detailed knowledge about the R&D portfolio and between these experts and their funders in Congress.

The four design principles and their components are summarized in Table 1.1.

**Table 1.1:** *Summary of Four Design Principles for an R&D Decision Making Process*

<b>Principle</b>	<b>Components of Principle</b>
1. <b>Quantifiable:</b> Technology improvement benefits prospectively quantified with a full account of uncertainty	<ul style="list-style-type: none"> <li>• Technology benefits estimated conditional on R&amp;D levels</li> <li>• Dependence between technological improvements modeled</li> <li>• Benefits over different time horizons considered</li> <li>• Uncertainty in technology benefits of R&amp;D modeled explicitly and estimated under common conditions</li> </ul>
2. <b>Comprehensive:</b> Social benefits evaluated in a common framework	<ul style="list-style-type: none"> <li>• At least one social benefit evaluated with common units</li> <li>• Dependence between R&amp;D benefits modeled</li> <li>• Accommodation for details of how technology improvement benefits were estimated</li> </ul>
3. <b>Adaptive:</b> Flexible to changing assumptions	<ul style="list-style-type: none"> <li>• Flexible to update for technological change</li> <li>• Flexible to update for policy changes</li> <li>• Capable of sensitivity analysis</li> </ul>
4. <b>Transparent:</b> Feasible transparency	<ul style="list-style-type: none"> <li>• Transparency of assumptions</li> <li>• Transparency of methods</li> </ul>

### 1.1.3 Proposed Methods in the Literature

Several studies in the literature on U.S. public energy R&D have proposed approaches to estimating the value of supporting energy technology R&D using a variety of analytical methods. Schock et al. (1999) and Nemet and Kammen (2007) estimate the appropriate level of energy R&D as the difference between the cost of meeting CO<sub>2</sub> emissions targets using assumptions of a business-as-usual (BAU) and an advanced technology scenario of costs. Davis and Owens (2003) use the concept of real options to estimate the value of investments in renewable energy R&D. And Blanford (2009) estimates the optimal allocation of R&D funds for renewable energy, nuclear energy, and



coal with CCS by defining two states for the cost of those technologies (BAU and low) and assuming that the probability of achieving the low-cost technology is an exponential function of R&D. All of these approaches, however, are based on assumptions that are not grounded in empirical data about the degree of technological innovation that will result from increased R&D investment. Other studies have used expert elicitations to determine the relationship between technology-specific public R&D and technology cost and performance in the United States and the European Union (Anadón et al., 2012; Baker et al., 2009a, 2009b; Baker and Keisler, 2011; Bosetti et al., 2012; Chan et al., 2011), but have not made the additional step of quantifying aggregate benefits of a portfolio of R&D investments. The National Research Council (2007a) proposed a method to evaluate individual DOE R&D programs using expert-assessed probabilities incorporated in a decision-tree framework to capture key uncertainties coupled with models to quantify benefits. However, this report does not propose a method to assess the aggregate benefits of an R&D portfolio in which the benefits of individual programs are contingent on each other. Nevertheless, the insights on assessing individual R&D programs were important for shaping this proposed approach to quantifying the benefits to energy R&D.

#### **1.1.4 The Challenges of R&D Benefit Estimation**

The uncertainties in the returns to R&D make estimation of the benefit of a single R&D program, let alone an entire R&D portfolio, challenging. Benefit calculation requires significant technological assumptions including an explicit representation of uncertainty. Uncertainty in the returns to R&D are well-documented, with a particular strain in the literature emphasizing the skewed distribution in the returns to R&D investment (Pakes, 1986; Scherer and Harhoff, 2000). It is also well-known that

uncertainty in the returns to R&D is a feature of the R&D resource allocation problem that makes it distinct from other investment problems (Arrow, 1962). This uncertainty can be due to: complementary and substitute technologies leading to inter-technology dependent or “recombinant” uncertainties (Fleming, 2001), the public goods nature of the information outcomes of R&D leading to spillovers between R&D investments by different (public/private) actors (David et al., 2000; Griliches, 1992; Keefer, 1991), and contingencies on many other factors (e.g., unpredictable demand, changing macroeconomic conditions, dependencies on markets in other countries, changing patterns of scientific/technical human capital etc.). Accounting for uncertainty in the returns to R&D is important for benefit estimation because greater uncertainty may either increase or decrease the value of conducting R&D (Bloom and Van Reenen, 2002; Santiago and Vakili, 2005). There is a considerable literature estimating the ex-post returns to R&D (see Hall et al. (2010) for a survey), with a particular strand grappling with the question of attributing technological change to specific R&D projects (Hall, 1996). However, there is a more limited literature on methods to quantify the ex-ante returns to R&D, which are precisely the quantity necessary to assess in designing an R&D portfolio.

Different methods in the literature can be classified in how they utilize information to quantify uncertainty ex-ante: (1) historic data (McNerney et al., 2011; Nemet, 2006; Wiesenthal et al., 2012); (2) data on early stage “precursor” technologies (Martino, 1987; Roper et al., 2011); and (3) data from technology experts. Methods that utilize data from technology experts have distinct advantages as an information source that feeds into a decision making process concerning R&D investments. Data from technology experts allows for the possibility that technologies may advance through new pathways that

endogenously depend on current decisions or in ways supported by only the most recent information. Further, technology experts can incorporate useful information that is unpublishable or proprietary.

Approaches to quantifying uncertainty typically represent probability distributions stylistically. Methods that utilize any of the three sources of information described can estimate notions of uncertainty parametrically (e.g. first and second moments) or non-parametrically (e.g. selected percentiles). Therefore, uncertainty analysis that requires propagating the uncertainty of estimated input parameters through a model typically depends on additional assumptions and appropriate sampling techniques. This is a particular challenge when there are multiple uncertain and non-independent parameters (such as the impact of R&D in various energy technologies that interact in the market), requiring an explicit representation of co-variation in parameters.

## 1.2 Methods

Inspired by an examination of the strengths and weaknesses of current practice in allocating U.S. public energy R&D investments, this chapter proposes a methodology to generate inputs to an R&D decision making process that satisfies the four proposed principles. This method has six components, detailed below and in Figure 1.5 (which uses public R&D investment in just two technology areas as an example).

This chapter develops and applies a method to support public energy R&D investments in the context of decision making at DOE. The first stage of this method is a major expert elicitation exercise that collected inputs from 100 experts in a range of sectors about the probabilistic distribution of the future cost and performance of 25 energy technologies conditional on different DOE R&D investment levels and allocations. Conditional probability distributions then parameterize the MARKET

ALlocation (MARKAL) model, a bottom-up energy system model of the U.S. economy with internal buy-in at DOE (as demonstrated by its use by the DOE Office of Policy and International Affairs in the past). The conditional probability distributions are introduced into MARKAL with a targeted sampling strategy to approximate joint distributions, allowing us to run Monte Carlo simulations to approximate the distribution of several outcome variables of interest, such as economic surplus, pollution emission levels, and crude oil imports. The final step of the methodology uses an importance sampling and optimization approach to estimate optimal R&D portfolios for varying aggregate budgets.

### **1.2.1 Expert Elicitation**

Recent research has shown that technical experts are able to consider and quantify the outcomes of R&D programs (Ruegg and Feller, 2003) and even parameterize probability distributions of outcomes (Chan et al., 2011; Catenacci et al., 2013; Bosetti et al., 2012; Fiorese et al., 2013; NRC, 2007a; Baker and Keisler, 2011; Baker et al., 2009b, 2009a; Curtright et al., 2008a; Abdulla et al., 2013; Jenni et al., 2013; Baker et al., 2010). Following the first design principle, the method proposed in this chapter uses expert elicitation to parameterize probability distributions of technological improvement returns conditional on several R&D levels. Data from technology experts is collected through an expert elicitation of over 100 experts in six technology areas (fossil energy, vehicles, energy storage, biofuels, solar energy, and nuclear energy)<sup>1</sup>, covering 25

---

<sup>1</sup> These elicitations were conducted as part of a broader study on the political economy of energy R&D investments in the United States. The full results of this study are published in Anadón et al. (2011) and Anadón et al. (2014a). A 7th elicitation on building technology was also conducted but not utilized in this chapter due to implementation complexities in MARKAL.

technologies<sup>2</sup>. R&D funding in these six technology areas totaled approximately \$1.5 billion in 2012, or approximately one-third of total DOE-funded applied R&D investment. Estimates were collected using six distinct written and online surveys administered between 2009 – 2011. Experts were selected for this study based on their contribution to the peer-reviewed literature, conference participation, and employment at research laboratories and universities. See Appendix A.1 for a more detailed description of the elicitation protocol and Anadón et al. (2014b) for even further detail, including the names of participants and the raw results of the elicitations.

Expert elicitation is a structured and systematic process for collecting and assessing subjective probabilistic estimates from individuals with particular expertise of interest (Anadón et al., 2012; Baker et al., 2009a; Chan et al., 2011; Cooke, 1991; Curtright et al., 2008b; Morgan, 2014; Morgan and Keith, 1995). This method involves in-depth engagement with experts to extract their subjective beliefs about uncertain parameters. This differs from surveys in that individual responses are not treated as observations from a single population, but rather as representative of a large body of knowledge. Therefore, expert elicitation seeks to include a group of participating experts of the highest quality and diversity of expertise, not the greatest quantity of viewpoints (Cooke, 1991). To avoid unwanted interactions between experts that would obscure the true diversity of judgments (Oppenheimer et al., 2007), experts were elicited individually

---

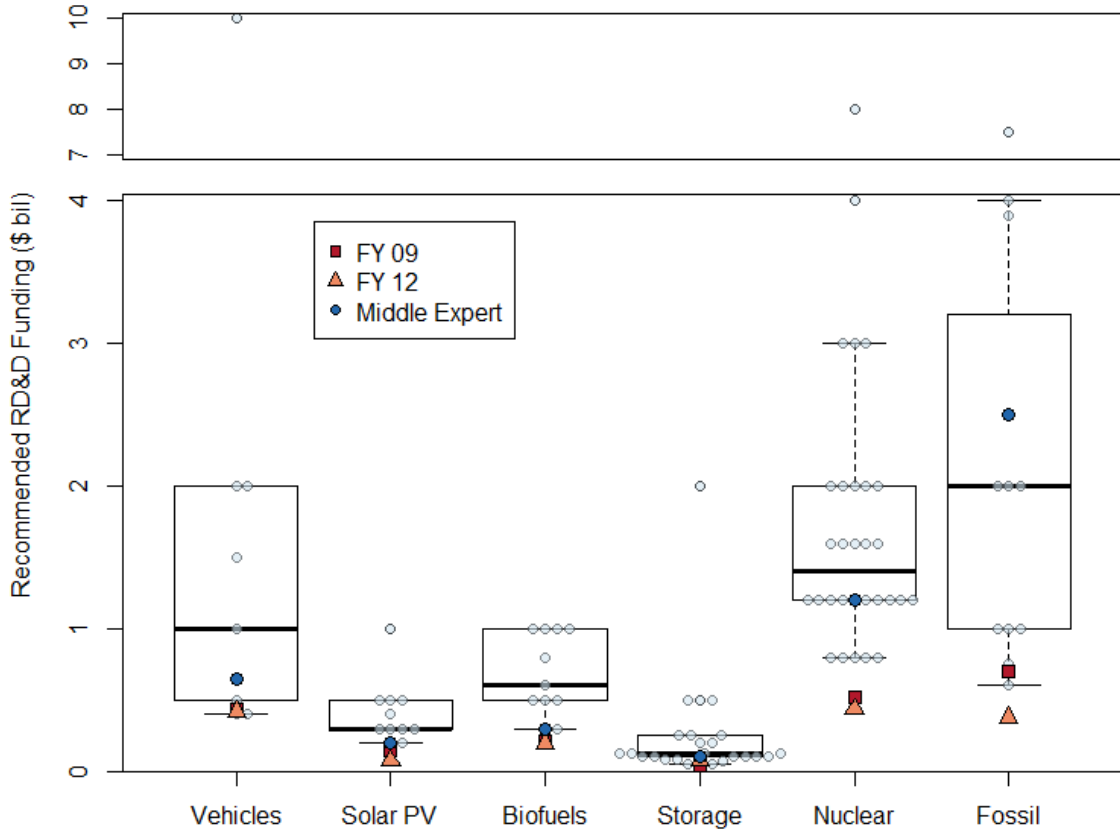
<sup>2</sup> Vehicles technology included fuel cell vehicles, battery-electric vehicles, plug-in hybrid vehicles, hybrid vehicles, and advance internal combustion vehicles. Solar photovoltaics (PV) included utility PV, residential PV, and commercial PV. Biofuels included bioenergy for electricity, and biochemical and thermochemical biofuels for jet fuel, diesel, and gasoline (thermochemical biofuels for jet fuel was excluded). Energy storage included grid-scale lithium ion batteries, sodium-sulfur batteries, flow batteries, flywheels, and compressed air energy storage. Nuclear technology included modular nuclear reactors, Generation IV reactors, and Generation III/III+ reactors. Fossil energy included natural gas power plants with and without carbon capture and coal power plants with and without carbon capture.

rather than in a group setting, as in the related Delphi process or expert consensus methods (Dalkey, 1969).

As in other technology forecasting expert elicitation, each elicitation instrument utilized in this study began with a technology primer based on a broad survey of the engineering literature in the technology area (Curtright et al., 2008a). These primers covered current technology cost and performance, fuel costs – if applicable, a summary of previous studies about future costs, and a summary of current U.S. federal government R&D investments in the particular technology area (primarily, but not exclusively, these investments were managed by DOE). Experts were then presented with an overview of heuristics to reduce their bias and overconfidence and asked to provide a detailed self-assessment of their expertise.

After this background material, experts were asked to provide their estimates of 2010 technology costs. Next, they were asked to recommend a level of R&D funding for the technology area that the study considered and propose a specific allocation of these funds to specific technologies and research pathways within the technology area. These recommendations were made without a consideration of the tradeoffs between technologies and without a presupposed policy goal (although experts were asked to describe their strategy for allocating funds to specific technology areas and along different technology development stages). Figure 1.1 displays the results of the R&D funding recommendation portion of the expert elicitation. Recommendations for R&D funding in the six technologies consisted of R&D in several sub-technologies and across different stages of development; for these specific results see Anadón et al., (2014b). From a methodological perspective, these results can be compared to the results from expert panels discussed above (American Energy Innovation Council, 2010; APS, 2008;

NCEP, 2004; PCAST, 2010, 1997). In every technology area, the majority of experts in this study recommended funding levels greater than current allocations.



The box plots are the expert-recommended funding levels (with empty grey circles denoting individual recommendations). The blue circles are the recommendations of the chosen “middle expert” (based on centrality of cost estimates, not recommendations). These middle expert recommendations form the basis of much of the subsequent analysis. The red and orange markers are the FY 2009 and FY 2012 DOE budget allocations to particular R&D areas, respectively.

**Figure 1.1:** *Comparison of Expert-Recommended R&D Levels*

In the next step of the elicitations, experts provided estimates of their subjective belief distribution for specific technology costs in 2030 under four R&D funding scenarios (business as usual and three hypothetical R&D scenarios based on multiples of their recommendations). The elicitation results are constrained by the way in which technical

experts are able to consistently understand and interpret probabilistic statements. Specifically, experts estimated their belief distributions by reporting the 10<sup>th</sup>, 50<sup>th</sup>, and 90<sup>th</sup> percentiles of the distribution of 2030 technology costs. Belief distributions were not elicited this way because percentiles can be expressed in the natural units of technology costs, which are more familiar to technology experts (unlike the inverse problem of estimating the cumulative probability of a given technology cost). Further, the inverse problem requires pre-specifying technology cost levels which could induce bias through anchoring or could be far from experts' beliefs and therefore not informative. Three percentiles were elicited to limit the number of parameters experts had to grapple with while still providing enough information to recover an estimated continuous distribution. This was a conscious decision in light of the limited patience and attention of participating experts and the tradeoff between more precisely estimating a smaller number of belief distributions or more imprecisely estimating a larger number of distributions. The elicitations concluded with a suite of qualitative questions to better understand each expert's R&D strategy.

### 1.2.2 Formalizing Elicitation Results

This subsection uses the following notation to describe the steps of how the expert elicitation outputs were translated into inputs for the MARKAL model:  $t$  indicates specific technologies which are grouped into technology clusters,  $c(t)$ ;  $p$  indicates the specific percentiles,  $\alpha_p$ , used in the expert elicitation;  $r$  indicates the R&D budget multipliers,  $\gamma_r$ , used in the expert elicitation;  $R_{c(t)}$  is the expert-recommended R&D levels for the  $c(t)$  technology cluster;  $RD$  are specific R&D funding levels;  $x_{t,p,r}$  is the cost of a technology  $t$  at the  $p$  percentile under the  $r$  R&D level. The  $c(t)$  technology clusters



group the 25 technologies into 6 R&D clusters based on how aggregate R&D investments are allocated to the specific technologies. For example, R&D investments in solar photovoltaic technologies are allocated to utility photovoltaics, residential photovoltaics, and commercial photovoltaics.

The outputs of the expert elicitations were: R&D budget recommendations for each technology-cluster ( $R_c$ ), technology cost estimates at three points on the inverse cumulative distribution (the 10th, 50th, and 90th percentiles) in 2010 and in 2030 at each of three multiples (0.5, 1, and 10) of recommended R&D levels. In 2030, the cost parameters,  $x_{t,p,r}$  for each of the  $t = 1, \dots, T$  ( $T = 25$ ) technologies were estimated by experts at three points on the inverse cumulative distribution function of cost:

$$x_{t,p,r} = Q_t(\alpha_p; RD_{c(t)} = \gamma_r R_{c(t)}) \quad (1.1)$$

In the above expression,  $c(t)$  is the R&D cluster of technology  $t$ ,  $RD_{c(t)}$  is the R&D level in the  $c(t)$  cluster,  $R_{c(t)}$  is the expert's recommendation for  $RD_{c(t)}$ ,  $p = 1,2,3$  index three percentiles ( $\alpha_1 = 0.1, \alpha_2 = 0.5, \alpha_3 = 0.9$ ), and  $r = 1,2,3$  index three R&D multipliers ( $\gamma_1 = 0.5, \gamma_2 = 1, \gamma_3 = 10$ ). 2010 costs were elicited as in Equation (1.1) and do not depend on R&D levels. With estimates for 2030 technology costs for twenty five technologies at three percentiles and three R&D levels, and estimates for 2010 technology costs at three percentiles, this amounts to a minimum of 300 total parameter estimates from the expert elicitation exercise.

An explicitly modeled dependence structure between these conditional distributions is represented as the  $V_T$  Spearman correlation matrix of 2030 costs, capturing co-variation in technology improvement benefits that could be due to technological spillovers (e.g. from grid-scale batteries for energy storage to batteries used in electric

vehicles). Spearman correlation, rather than Pearson correlation, is appealing for expert elicitation because it is cognitively more natural to think of relationships between uncertain quantities in terms of their strength of monotonicity rather than their strength of linearity. For many technology-technology diodes, a correlation in the  $V_T$  matrix was not elicited and instead it was assumed that technology costs would be independent.

Similarly,  $T$  estimates of  $\rho_t$ , are estimated time rank correlations for each technology. The  $\rho_t$  time correlations take into consideration how much information relative 2010 technology costs provide for relative 2030 technology costs. These longitudinal (rank) correlations will be high for most technologies ( $0.7 \leq \rho \leq 0.9$ ), as experts assumed that it is unlikely that a technology on the high end of costs in 2010 will be on the low end of costs in 2030.

In the vehicle technologies elicitation instrument, experts were asked to provide information that could be used to empirically estimate joint probability distributions of costs that were related to batteries. After asking experts to provide their 10<sup>th</sup>, 50<sup>th</sup>, and 90<sup>th</sup> percentile cost estimates in 2030 under the various R&D scenarios, experts were asked to assume as fixed a given cost of a technology in 2030 and then re-estimate the distribution of costs for a different technology in 2030. Through this conditional re-estimation, experts provided information that could be used to estimate a correlation between the future costs of the two technologies. However, these suites of questions were not implemented in all surveys because these were mentally taxing questions on an already fatiguing exercise. Therefore, a Delphi process within the research group was used, which had significant technology expertise, to propose a correlation matrix of 2030

technology costs for all twenty-five technologies. For many technology-technology diodes, it was assumed that technology costs would be independent (See the Appendix A.2 for the correlation matrix used in this work).

The elicitations utilized included subjective belief distribution estimates from over 100 experts. However, for analytical tractability and the well-known problems of inter-expert reliability’s dependence on multiple factors (Anadón et al., 2013; Keith, 1996), a representative “middle” expert for each of the six elicitations was selected for the results presented in this work. Other analysis not focused on the methodological aspects could analyze the results of using middle experts alongside with the results of using “optimistic” and “pessimistic” experts to further explore the impact of expert selection on outcomes. “Middle” experts for each of the 6 technology areas were selected by evaluating which expert in each area had central estimate,  $x_{t,2}$ , and uncertainty range indicator,  $(x_{t,3} - x_{t,1})/x_{t,2}$ , estimates of future technology costs that fell at or near the average values of central and uncertainty range estimates of all experts in their area. Selection of “middle” experts was made without considering the expert’s R&D funding recommendation (see Figure 1.1). This selection was also vetted with 23 “higher level” qualitative reviewers with experience in the management of large scale technology programs (Anadón et al., 2014b).

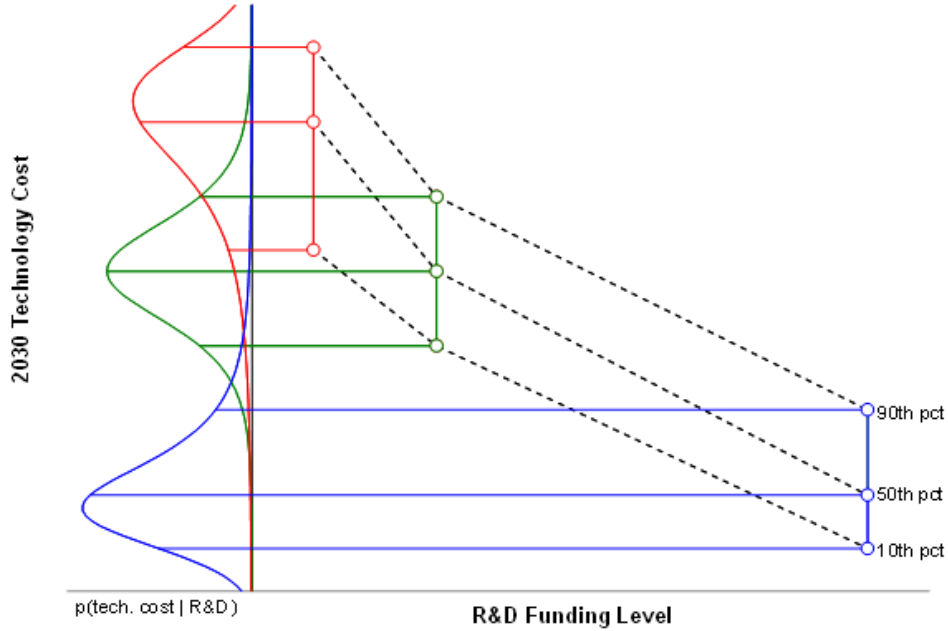
### **1.2.3 Sampling Strategy**

Using the expert elicitation results from middle experts, estimates of the benefits of an R&D portfolio are generated by introducing these estimates in the MARKAL model, a publicly-available energy-economic model (see Section 1.2.4 for detail of the model). A key challenge is to translate the expert elicitation results into congruent probability

distributions that could efficiently estimate outcome distributions given the computational cost of MARKAL. Because MARKAL is so computationally expensive due to the level of technical detail it incorporates, this sampling strategy is constrained to a finite number of model simulations (1,200 scenarios in the final implementation used in optimization). To overcome the limits of this constraint, samples from the distributions of technology costs were drawn using Latin Hypercube sampling (LHS) (McKay et al., 1979) and in a way to cover the full range of possible R&D realizations. LHS reduces the total variance in the sampled values of uncertain quantities by fixing the probability mass between all sampled values but does not introduce bias. The variance reduction property of LHS allows us to decrease the total number of samples (and therefore total number of model simulations) that are needed to run in order to sufficiently describe the variance in the output quantities of interest. However, the tradeoff in using LHS is that the samples drawn from the process are no longer independent and therefore, each sample viewed by itself is difficult to interpret.

A shifted-log-logistic (SLL) functional form is imposed to interpolate and extrapolate the expert elicited percentiles as full probability distributions. The SLL distribution has three parameters, allowing us to exactly identify a continuous probability distribution that maximizes the information content of the elicitation results. Equation (1.2) shows the quantile function parameterization of the SLL distribution conditional on  $\mathbf{x}_t$  in terms of the three values of  $x_{t,p}; RD_{c(t)}$  at a given level of R&D, allowing us to sample values of  $p$ . A graphical depiction of the SLL fitting is shown in Figure 1.2 for percentiles estimated by experts at three R&D levels.

$$Q_t(p|\mathbf{x}_t; RD_{c(t)}) = \frac{(x_{t,3} - x_{t,2})(x_{t,1} - x_{t,2})}{2x_{t,2} - x_{t,3} - x_{t,1}} \left[ \left( \frac{1-p}{p} \right)^{\log \left[ 1 + \frac{2x_{t,2} - x_{t,3} - x_{t,1}}{\log(9)(x_{t,3} - x_{t,2})} \right]} - 1 \right] + x_{t,2} \quad (1.2)$$



A graphical depiction of representative sets of expert-elicited percentiles for three R&D scenarios (shown separately in red, green, and blue white-filled circles) and conditional probability distribution fitting with the SLL distribution (on the left side of the y-axis).

**Figure 1.2:** *Distribution-Fitting to Expert-Elicited Percentiles*

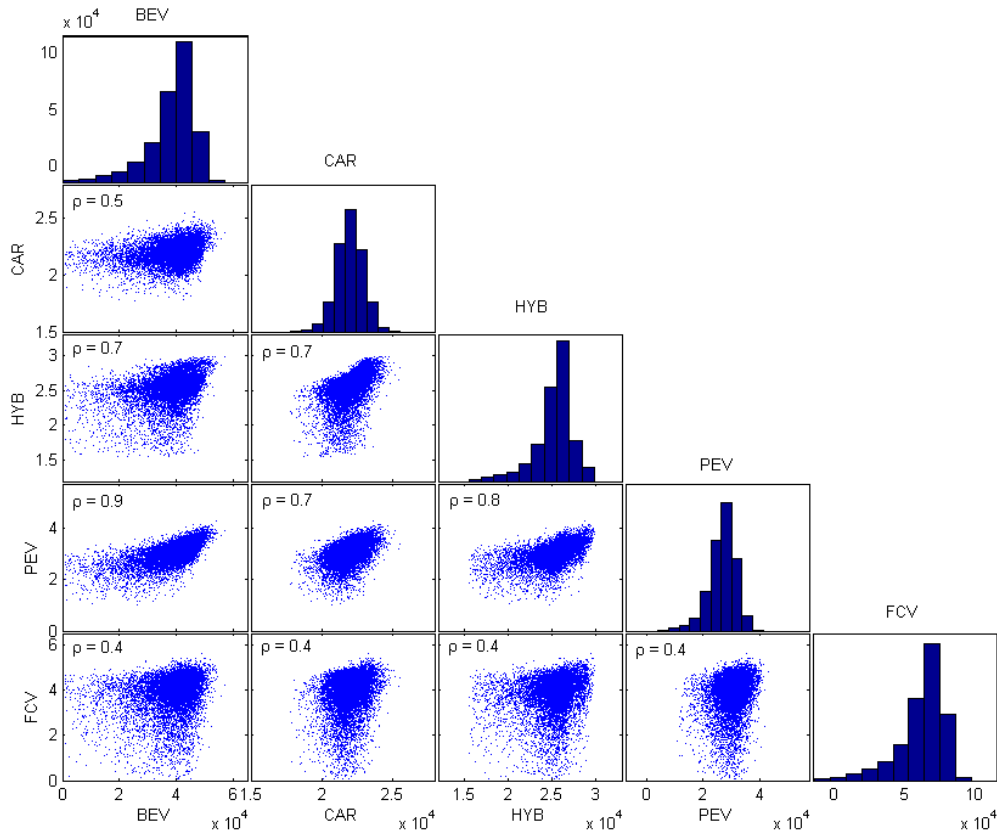
While there are many probability distributions with three parameters, the SLL distribution is relevant for this application because it is a smooth distribution that can allow for skewness. A disadvantage of the SLL distribution is that it is almost always bounded on one side and the bounds are determined by the three parameters. While there may be conceptual reasons to include bounds in the distribution of costs (e.g. forcing costs to be greater than zero), bounds could also be elicited directly, although this would add to the number of parameters included in elicitations. Ultimately, the choice of

any probability distribution that can fit the three percentiles exactly is arbitrary since a three parameter distribution fully utilizes the information content of the elicited beliefs.

Our elicitation methodology included a process for estimating a Spearman correlation matrix of 2030 technology costs,  $\mathbf{V}_T$ . Using the  $T$ -dimensional Gaussian copula,  $C_{\mathbf{V}_T}^{Gauss}$  for each sample,  $T$  samples from Unif(0,1) with the specified  $\mathbf{V}_T$  Spearman correlation matrix are drawn. Then, correlated samples from the joint distribution of cost given a fixed level of R&D can be found by evaluating the quantile function given in Equation (1.2) at the correlated variates using  $(\mathbf{x}_t; RD_{c(t)})$  as calculated in Equation (1.1). More formally, joint samples of 2030 technology costs,  $\mathbf{x}^*$ , are drawn from a joint distribution such that the variates are distributed as described in Equation (1.3).

$$F(\mathbf{x}^*|\mathbf{RD}) \sim C_{\mathbf{V}_T}^{Gauss} \quad (1.3)$$

Because  $(x_{t,p}; RD_{c(t)})$  is weakly monotonic in  $RD_{c(t)}$ , the rank correlations as sampled from the copula are maintained under this transformation. An alternative to this approach proposed by Iman and Conover (1982) essentially approximates the Gaussian Copula with less precision. Figure 1.3 shows example joint samples for 5 vehicle technologies with correlated costs.



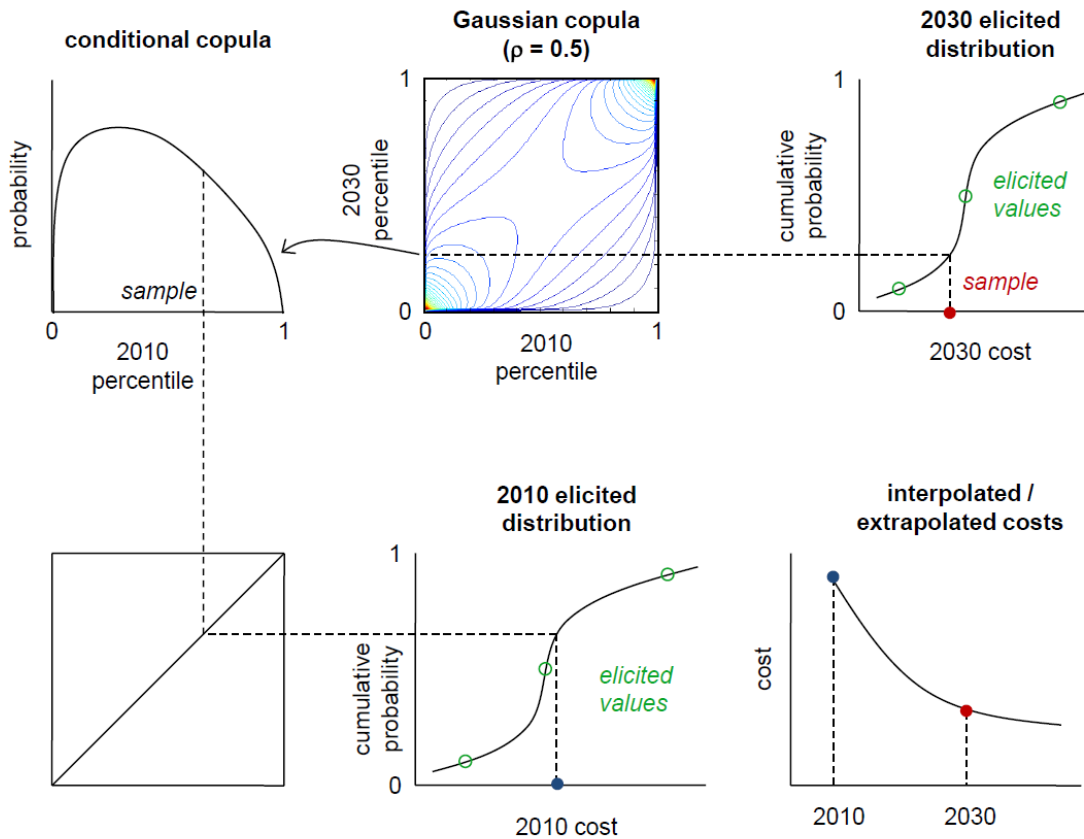
An example sample from the joint distribution of 2030 costs for five technologies in 2010 USD: battery electric vehicles (BEV), advanced internal combustion vehicles (CAR), hybrid vehicles (HYB), plug-in electric vehicles (PEV), and fuel cell vehicles (FCV). Marginal sample distributions are shown as histograms in the diagonal plots and joint samples between all pairs are shown in scatter plots off the diagonals along with the corresponding rank correlation coefficients,  $\rho$ .

**Figure 1.3:** *An Example Sample from the Joint Distribution of Five Technology Costs*

Next, for each sample the 2010 and 2030 cost draws are interpolated and extrapolated to construct a vector of costs for 2010–2050 in 5-year time steps. This step is necessary to conform to the required inputs of the MARKAL model. Because the elicitations only provide estimates of costs in two years (2010 and 2030), additional assumptions are required, which have a substantial effect on the 2030–2050 extrapolation region. With the 2030 costs sampled, 2010 costs conditional on 2030 costs are sampled using the estimated time correlation,  $\rho_t$  and the expert-elicited probability

distribution for 2010 costs, fit in the same way as the 2030 distribution. To calculate a sample's 2010 cost, a 2-dimensional copula with rank correlation  $\rho_t$  is sampled, conditioning on the 2030 sample's quantile (which was sampled from a  $T$ -dimensional copula). This approach samples from the 2-dimensional copula, representing time-dependence, conditionally using a numerical approximation at  $10^{-5}$  precision. Then, the resulting quantiles for 2010 are transformed by Equation (1.2), using the  $x_t$  values elicited for technology  $t$  in 2010. Note that for 2010 estimates, there is no need to condition on R&D levels since R&D is constant by definition in 2010. For each sample, if the 2030 draw is greater than the 2010 draw, costs are linearly interpolated and extrapolated; otherwise, an exponential functional form is imposed. The choice of functional form is of little practical consequence in the interpolation region (2010–2030), but has important consequences for the extrapolation region (2031–2050). The choice of a linear functional form for scenarios with cost increases is consistent with cost increases due to institutional causes, such as increasing permitting and licensing costs, which are unlikely to evolve at nonlinear rates. The choice of an exponential functional form for cost decreases is conservative in the sense that the rate of cost decrease declines from 2030–2050, relative to the rate of cost decrease from 2010–2030. See Figure 1.4 for a graphical depiction of these steps.





A graphical depiction of dependent sampling of 2010 costs conditional on a sampled 2030 cost. Moving from the top right to the left, the 2030 cost is sampled (shown in red) from the estimated distribution based on elicited percentiles (in green). In the top center figure, the 2-dimensional Gaussian copula at the 2030 sampled cost defines a conditional probability distribution of 2010 cost percentiles, shown in the top left subfigure. The bottom left figure reflects the 2010 cost percentiles, which are then transformed across the 2010 elicited distribution, shown in the center frame of the bottom row. These two samples together are then used to interpolate costs between 2010 and 2030 using a linear functional form for scenarios with cost increases and an exponential functional form for cost decreases. To extrapolate costs to 2050 it is assumed that between 2030 and 2050 costs follow the same functional form as between 2010 and 2030.

**Figure 1.4:** *Graphical Depiction of Dependent Sampling*

## 1.2.4 Modeling with MARKAL

The method proposed in this chapter parameterizes the results from the elicitations to use as stochastic cost inputs in MARKAL, a detailed energy system model (Fishbone and Abilock, 1981; Loulou et al., n.d.). This approach also uses elicited point values of performance metrics, which were held constant across samples at expert-elicited values

but varied over time. Other model parameters not included in the elicitations were held constant at their default values based on the U.S. Energy Information Administration's Annual Energy Outlook (EIA, 2009). MARKAL is a bottom-up, partial equilibrium model of the U.S. economy that is specifically designed to represent technological evolutions of the physical energy system occurring over 30- to 50-year periods. MARKAL is solved as a cost minimization problem where future states of the energy system are determined by identifying the most cost-effective pattern of resource use and technology deployment over time, given exogenously specified energy demands (Anadón et al., 2014b; Fishbone and Abilock, 1981; Loulou et al., n.d.). DOE and EPA have each developed their own versions of MARKAL for their in-house policy analysis. This study utilizes a version of the U.S. multi-region MARKAL model maintained by Brookhaven National Laboratory, one of the main operators of MARKAL for DOE.

MARKAL was chosen for this study for its technical detail which allows us to accommodate the nuances of the impact of the technology improvement benefit estimates generated in the expert elicitations. For example, MARKAL allows us to include the specific performance characteristics of technologies that the experts conditioned on when estimating costs, the solar irradiation profile and underground CO<sub>2</sub> storage space in each of MARKAL's 10 geographic regions, and the interaction between different vehicle types in satisfying aggregate vehicle demand. Further, MARKAL allows us to evaluate social benefits along a common set of metrics while also accounting for the interactions of technologies in satisfying market demand.

The MARKAL model allows us to integrate the results from the suite of expert elicitations in different technology areas in a transparent framework with its own

assumptions that can be individually assessed<sup>3</sup>; MARKAL is publically-available and many government agencies implement the model themselves. A MARKAL scenario run produces hundreds of outcome metrics of potential interest (e.g., CO<sub>2</sub> emissions, energy costs, oil imports, etc.), many of which could be used to evaluate an R&D portfolio.

### 1.2.5 Optimization

Our approach implements a sampling and optimization method that allows for the estimation of the optimal allocation of R&D investments at a range of budget levels. For these results, 1,200 MARKAL Monte Carlo samples of technology costs are run under a wide range of R&D levels that cover the full range of R&D scenarios considered. An importance sampling technique is applied that allows for the calculation of the expected value of outcome metrics under specific R&D portfolios that are not pre-specified (Morgan and Henrion, 1998). This importance sampling approach allows for the flexibility to update the technical assumptions of the study without repeating the prior steps of the method, satisfying the third design principle. The importance sampling strategy allows one to readily adjust for changing input assumptions—such as different R&D levels in the different technologies in the portfolio—without requiring additional model runs, thus solving a computational constraint faced by many decision-making entities that would otherwise only be able to evaluate a small number of proposed R&D portfolios (Pugh et al., 2011). Additionally, this method’s flexibility to different input assumptions may be particularly useful in testing the potentially important sensitivity of

---

<sup>3</sup>The advantages of using a credible and transparent model in a decision making process that involves technical knowledge held by interested parties is notably discussed in the context of the Long Range Transport of Atmospheric Pollutants (LRTAP) Protocol and the use of the RAINS model during the negotiations (Figueira et al., 2005; Parson, 2002, 1998). By agreeing on methodology prior to the introduction of technical assumptions, parties built credibility.

results to assumptions from different sources (e.g., more optimistic experts, experts internal to the decision making process, experts from stakeholder groups, or experts from different countries) and elicitation strategies (e.g., in person interviews, or written surveys) (Anadón et al., 2013).

Using this importance sampling strategy, a response surface of the expected returns across a continuous range of R&D levels can be constructed—as represented in subplot F of Figure 1.5. The final step of this method is to fit a polynomial response surface to the expected outcomes from the importance sampler and use a numerical optimization algorithm to calculate the R&D portfolio allocation that yields the optimal outcome (e.g. highest economic surplus) for a fixed R&D budget. This approach optimizes the portfolio on the expectation of the outcome metric. This method follows Principle 4 by using a publicly available economic model and explicitly representing conditional distributions of the returns to R&D in a way that could be easily communicated to actors external to the decision making process. The names of the experts that contributed to the estimates of the impact of public R&D on future technology cost and performance are also public (Anadón et al., 2011).

Our approach inputs stochastic realizations of technology cost parameters conditional on randomly drawn R&D levels and holds all other model parameters constant across model runs at their default values (i.e. technology performance parameters, which were held constant across samples at expert-elicited values but varied over time, and other model parameters not included in the elicitations, which were held constant at their default values). While the MARKAL model outputs many metrics for a given sampled cost vector, this method uses only one outcome metric in the optimization stage, denoted  $S(\mathbf{x}^*)$ .

The first stage of the optimization is to use the relationship between R&D funding and technology costs from the expert elicitation and the relationship between technology costs and the outcome metric,  $S(\mathbf{x}^*)$ , from many samples, to estimate the expected value of the outcome metric at any specified R&D level in the feasible range. Formally, for a given R&D vector,  $\mathbf{RD}$ , the goal is to find the expression in Equation (1.4).

$$\mathbb{E}[S(\mathbf{x}^*)|\mathbf{RD}] = \int S(\mathbf{x}^*)p(\mathbf{x}^*|\mathbf{RD}) d(\mathbf{x}^*|\mathbf{RD}) \quad (1.4)$$

However, the distribution  $p(\mathbf{x}^*|\mathbf{RD})$  is computationally intractable for integration due to the complexity of the expert-elicited distributions. While this distribution could be simplified to allow for a more direct evaluation of the expectation in Equation (1.4), doing so would undermine the integrity of the expert assessments. Instead, this approach uses an importance sampling strategy to evaluate the expected value in Equation (1.4), using  $p(\mathbf{x}^*|\mathbf{RD})$  as the “target distribution.” The importance sampler is represented in Equation (1.5) where  $\gamma$  indexes the  $n=1,200$  MARKAL model runs that computational constraints allow, and  $k(\mathbf{x}_\gamma^*)$  is a kernel approximation to the entire joint “sampling distribution” of 2030 costs, described in Equation (1.3). The kernel approximation is used to calculate the “sampling probability” to reduce computational requirements and is of little practical consequence.

$$\mathbb{E}[S(\mathbf{x}^*)|\mathbf{RD}] \approx \frac{1}{n} \sum_{\gamma} \frac{p(\mathbf{x}_\gamma^*|\mathbf{RD})}{k(\mathbf{x}_\gamma^*)} S(\mathbf{x}_\gamma^*) \quad (1.5)$$

In the second stage of the optimization, the importance sampling strategy is applied to a grid of R&D vectors that span the feasible R&D space to calculate the expected outcome metric over the full range of possible R&D portfolios. An 8-unit grid in the 6-dimensional R&D space is used, yielding importance sampling calculations at  $8^6 = 262,144$  R&D vectors. The grid constructed evaluates all permutations of R&D levels at

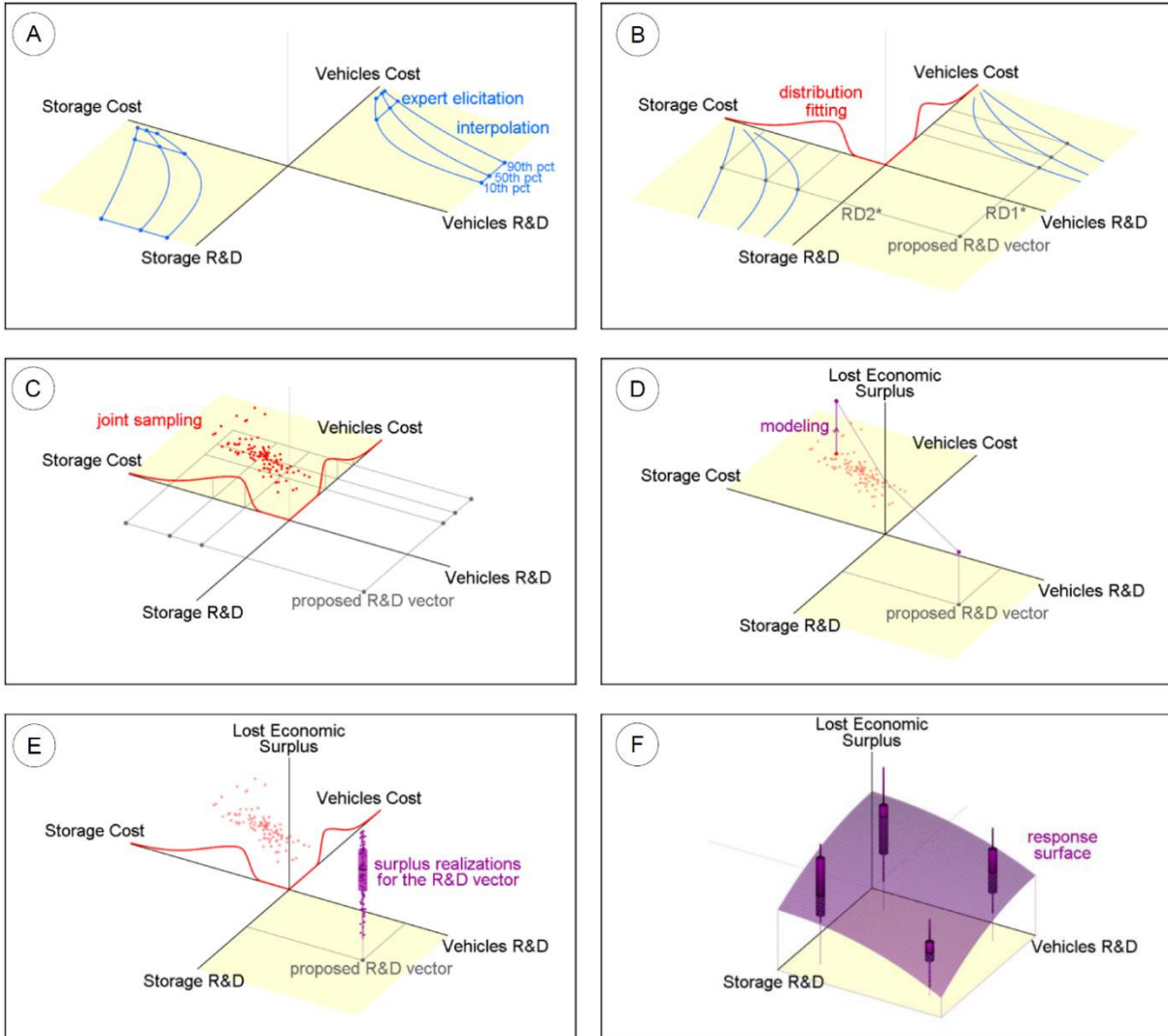
$\gamma = 0.5, 0.75, 1, 1.5, 2, 5, 7.5,$  and 10 multiples of the expert-recommended R&D level. These levels were chosen to give higher resolution to R&D levels close to the recommended levels while still informing higher R&D levels. Repeating the importance sampling strategy is the most computationally expensive step of the method. Future work could address this by investigating techniques to reduce the computational burden of evaluating the importance sampler's sampling distribution.

In the final stage of the optimization, a high dimensional polynomial is fit to the grid of expected outcome metrics. In the results presented in the chapter a least-squares fit is used to find a “response surface” of the expected outcome metric using the six first-order R&D terms, the thirty six second-order R&D terms (including squared terms and interactions). Since the polynomial fits the expected values of the outcome metric, the predicted values along the response surface can be thought of as a double expected value or a predicted value of an expectation, as shown in Equation (1.6).

$$\mathbb{E}[S|\mathbf{RD}] = \hat{\beta}_0 + \sum_c^6 \hat{\beta}_c RD_c + \sum_c^6 \sum_{c'}^6 \hat{\beta}_{c,c'} RD_c RD_{c'} \quad (1.6)$$

The estimated polynomial in Equation (1.6) is used in an optimization scheme to find the vector  $\mathbf{RD}_\Omega^*$  that maximizes  $\mathbb{E}[S|\mathbf{RD}]$  under the constraints  $\sum \mathbf{RD} \leq \Omega$ , where  $\Omega$  is a budget constraint, and that all R&D levels are within their feasible ranges. The results of repeating this optimization at many levels of  $\Omega$  are shown in Figure 1.7.

Figure 1.5 summarizes the six steps of this method using an example of two technology areas. The figure shows the collection of expert belief distributions, fitting probability distributions, sampling from the joint distribution, modeling with MARKAL, the Monte Carlo interpretation of modeled results, and finally the response surface connecting the set of Monte Carlo results.



This figure utilizes projections across relevant dimensions to visualize a schematic example of the methodology described in this chapter using two technologies as examples, energy storage and vehicles. Subplots (A) and (B) replicate the key features of Figure 1.2, namely probability distribution fitting to the expert-elicited cost percentiles at a given level of R&D. The expert elicitation provided estimates of the 10th, 50th, and 90th percentile of technology costs in 2030 under three R&D scenarios – these are shown with the blue points in the two R&D funding–technology cost spaces highlighted in yellow. For a given vector of R&D funding levels in storage and vehicle technologies ( $RD1^*, RD2^*$ ), shown by the grey line perpendicular to the R&D axes, the approach fits the three-parameter SLL probability distribution to the interpolated percentiles for technologies, shown in red. In subplot (C), then the joint distribution of technology costs is sampled, utilizing a dependence structure for technology costs – samples are shown as red dots in the highlighted storage cost–vehicle cost space; this replicates the key features of Figure 1.3. In subplot (D), for a vector (pair) of technology costs, aggregate outcomes, such as lost economic surplus, are estimated using the MARKAL model—the purple line perpendicular to the storage cost–vehicle cost plane shows this relationship. In subplot (E), using a LHS approach over storage cost–vehicle cost combinations, the distribution of outcome metrics is estimated, as shown with the purple vertical dot plot and superimposed box plot. The boxplot replicates the key features of Figure 1.6. Finally, in subplot (F) a response surface is built connecting the box plots that can be used to find the portfolio that optimizes the expected outcome metric under a budget constraint.

**Figure 1.5:** Summary of Methodology from Elicitations to an Optimal Portfolio

## 1.3 Results

This section evaluates the results of the suite of model runs parameterized by the expert elicitations in several ways. This section considers, in turn, system benefits (e.g. aggregate economic surplus and CO<sub>2</sub> emissions), technology diffusion (e.g. deployment of renewable energy), and uncertainty analysis (e.g. high and low percentiles of outcome metrics and the variance in outcome metrics). Each of these three classes of methods can be useful inputs to support public energy R&D investment decision making. These results are summarized in Figure 1.6, which focuses on a comparison of a BAU R&D portfolio that perpetuates Fiscal Year 2009 R&D investment allocations and levels at \$2.1 billion<sup>4</sup> and the “middle” expert-recommended R&D portfolio with a total investment of \$5.3 billion<sup>5</sup>.

### 1.3.1 System Benefits

In Figure 1.6, the top subplot shows that the recommended R&D portfolio reduces the median projection of annual CO<sub>2</sub> emissions by 46 million metric tons relative to the BAU portfolio in 2030 and by 253 million metric tons in 2050. Even with the recommended R&D portfolio, without additional limits, results project that CO<sub>2</sub> emissions will rise by 7% between 2010 and 2020, inconsistent with the stated goal of President Obama (Executive Office of the President, 2013) of a 17% reduction of total greenhouse gas emissions below 2005 levels by 2020.

---

<sup>4</sup> In FY 2009, DOE’s R&D portfolio in the six technology areas considered allocated \$214 million to bioenergy, \$432 million to vehicles, \$172 million to solar PV, \$701 million to fossil energy, \$514 million to nuclear energy, and \$83 million to energy storage (Gallagher and Anadón, 2014).

<sup>5</sup> The middle expert-recommended R&D portfolio allocates \$300 million to bioenergy, \$650 million to vehicles, \$200 million to solar PV, \$2,850 million to fossil energy, \$1,200 million to nuclear energy, and \$100 million to energy storage.



In terms of other environmental performance metrics, median projected NO<sub>x</sub> emissions under the recommended R&D scenario are 2% lower relative to the BAU scenario in 2030 and 5% lower in 2050. SO<sub>2</sub> emissions show smaller differences – in 2030 they are virtually equivalent, and in 2050, the recommended scenario has median projected emissions 3% lower than the BAU scenario.

The results in the middle subplot of the figure show that the median projection of annual economic surplus<sup>6</sup> is \$29 billion higher in 2030 and \$54 billion higher in 2050 with the recommended R&D portfolio relative to the BAU portfolio. Given that the recommended R&D portfolio has a budget \$3.2 billion per year greater than the BAU portfolio, the recommended portfolio has positive net social benefits.

### **1.3.2 Technology Diffusion**

The results in the bottom subplot of Figure 1.6 show that median renewable energy deployment (energy generated by hydroelectric, wind, solar, and biomass) under business as usual R&D is expected to increase from 2010 levels by 160% by 2030 and by over 230% by 2050. Under the recommended R&D scenario, these growth rates increase to 170% and 250%, respectively. In levels, the BAU R&D scenario increases renewable energy deployment from 2010 levels (approximately 485 TWh) to 770 TWh in 2030 and 1,130 TWh in 2050, while the recommended portfolio increases renewable energy deployment to 810 TWh and 1,190 TWh, respectively. This shows that secular trends in renewable energy deployment are much larger than the incremental deployment induced by the recommended additional R&D. However, these estimates only represent changes

---

<sup>6</sup> Net (producer and consumer) economic surplus is estimated by MARKAL as the area between the supply and demand curves for the full set of goods in the model and does not include most environmental externalities, such as greenhouse gas emissions.

due to R&D in a subset of renewable technologies since other potentially important renewable energy sources, such as geothermal, solar thermal, and wind, are not considered in the elicitation.

Unlike for renewable technologies, R&D in coal technology is unlikely to substantially alter projected deployment rates from 2010-2030 (perhaps due to the long-term nature of coal power plant construction and the other institutional factors that build inertia into the system of coal utilization). Projections through 2030 for coal deployment are virtually identical under the BAU R&D scenario and the recommended R&D scenario, with both scenarios estimating a median growth rate between 2010-2030 in coal energy of 22-24%<sup>7</sup> (10.7-10.9 EJ in 2030 relative to 8.8 EJ in 2010). Taken together, these results indicate that R&D policy alone is unlikely to substantially affect the deployment of coal energy.

For crude oil imports, as with coal, high inertia in the system limits the ability of R&D to affect outcomes in the short run. In 2030, the recommended R&D scenario has median projected oil imports 2% lower than the BAU scenario. However, in 2050, this number increases to 10%. While median projected crude oil imports are projected to decrease from 2010 to 2035, even under the BAU scenario, they are still projected to increase from 2035 to 2050. Median crude oil price projects are within 2% of each other under the two R&D scenarios from 2010-2050. Taken together, this suggests that other non-R&D policies would likely be needed to perpetually reduce net oil imports.

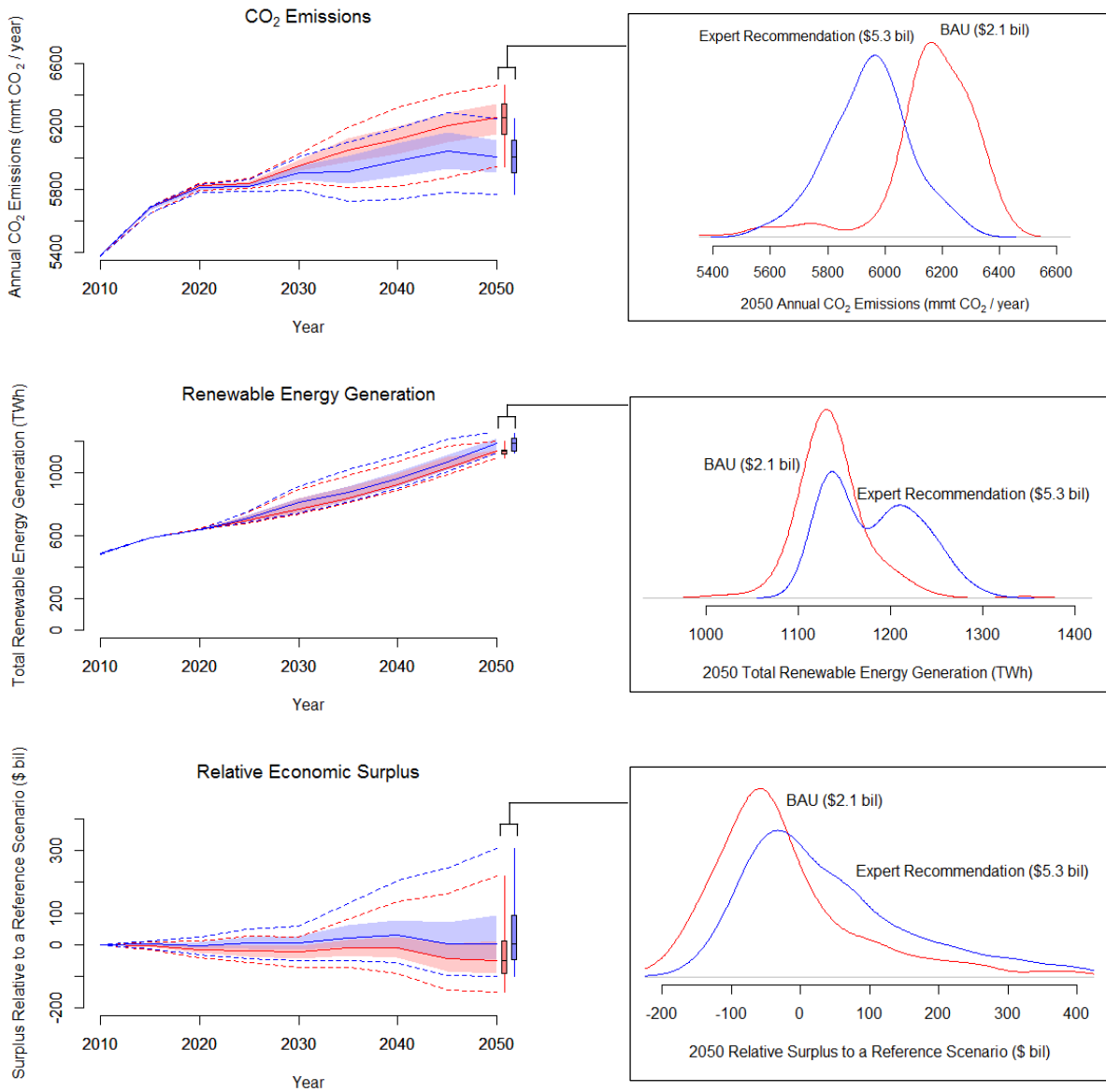
---

<sup>7</sup> Note that the model and expert elicitation were conducted during a period before natural gas prices were forecasted to decline as rapidly as they have in recent years due to large shale gas production.

### 1.3.3 Uncertainty Analysis

Figure 1.6 highlights the ability of this method to quantify the uncertainty in evaluation criteria of R&D portfolios. The figure allows the estimated distribution of evaluation criteria in different individual R&D portfolios to be assessed and compared as part of the R&D portfolio decision making process (DOE External Expert Peer Review Panel, 2006). For example, the proposed method allows for the quantification of high and low percentiles of outcome metrics. Relative to the business as usual R&D scenario, if the 5<sup>th</sup> percentile or 95<sup>th</sup> percentile of economic surplus is realized, the benefits of the recommended R&D scenario could be \$48.2 billion and \$89.7, respectively. As another example, under the BAU R&D scenario, the difference between the projected 95<sup>th</sup> percentile and 5<sup>th</sup> percentile of CO<sub>2</sub> emissions in 2050 is as large as 60% of the difference in median projected CO<sub>2</sub> emissions for 2050 and 2010.

Uncertainty can also be evaluated by looking at the variance in outcome metrics. These results show that the variance in projected economic surplus in 2030 and in 2050 is statistically greater with the recommended R&D portfolio than in the BAU case (F-test for difference in variances has p-value = 0.001 for 2030 surplus projections and 0.01 for 2050 surplus projections). This demonstrates that while greater R&D has positive expected benefits, additional R&D also creates more uncertainty in outcomes.



Estimation of benefits of individual R&D portfolios in terms of CO<sub>2</sub> emissions, relative economic surplus, and renewable energy generation. In the left figures, the dotted lines encapsulate the estimated 90% probability intervals (5<sup>th</sup> – 95<sup>th</sup> percentiles) and the lightly-shaded regions are estimated 50% probability intervals (25<sup>th</sup> – 75<sup>th</sup> percentiles) from 400 Monte Carlo samples using the “middle” experts for each of the 6 technology areas. The blue lines/regions show projections under the business as usual R&D funding portfolio, whereas the red lines/regions show projections under the expert-recommended R&D portfolio. In the right plots, kernel density approximations to the distribution of benefits in 2050 are shown under the same two R&D scenarios. The top subplot shows CO<sub>2</sub> emission benefits; the middle subplot shows economic surplus benefits, and the bottom subplot shows renewable energy generation.

**Figure 1.6:** *Distribution of Benefits under Two R&D Scenarios*

Greater expected benefits combined with greater uncertainty under greater R&D investments also translates into a quantifiable larger probability of desirable outcomes. For example, this analysis can be used to estimate that under the BAU R&D scenario, the probability that CO<sub>2</sub> emissions in 2050 will be below 6 Gt-CO<sub>2</sub> is 9%, while under the recommended scenario the probability would increase to 66%. The use of probabilistic language to talk about the uncertain benefits of R&D through a combination of expert elicitations and models using a transparent process would be a positive development in the decisions about budget allocations, as it would help move debates away from the credibility of final estimates towards the technical assumptions that matter, which are usually less politically vulnerable.

#### **1.3.4 Optimized R&D Portfolios**

This section presents an analysis of optimal R&D portfolio allocation under a no-policy scenario, shown in Figure 1.7. Overall, results yield four main insights about the optimal allocation of R&D resources across the six technology areas that investigated.

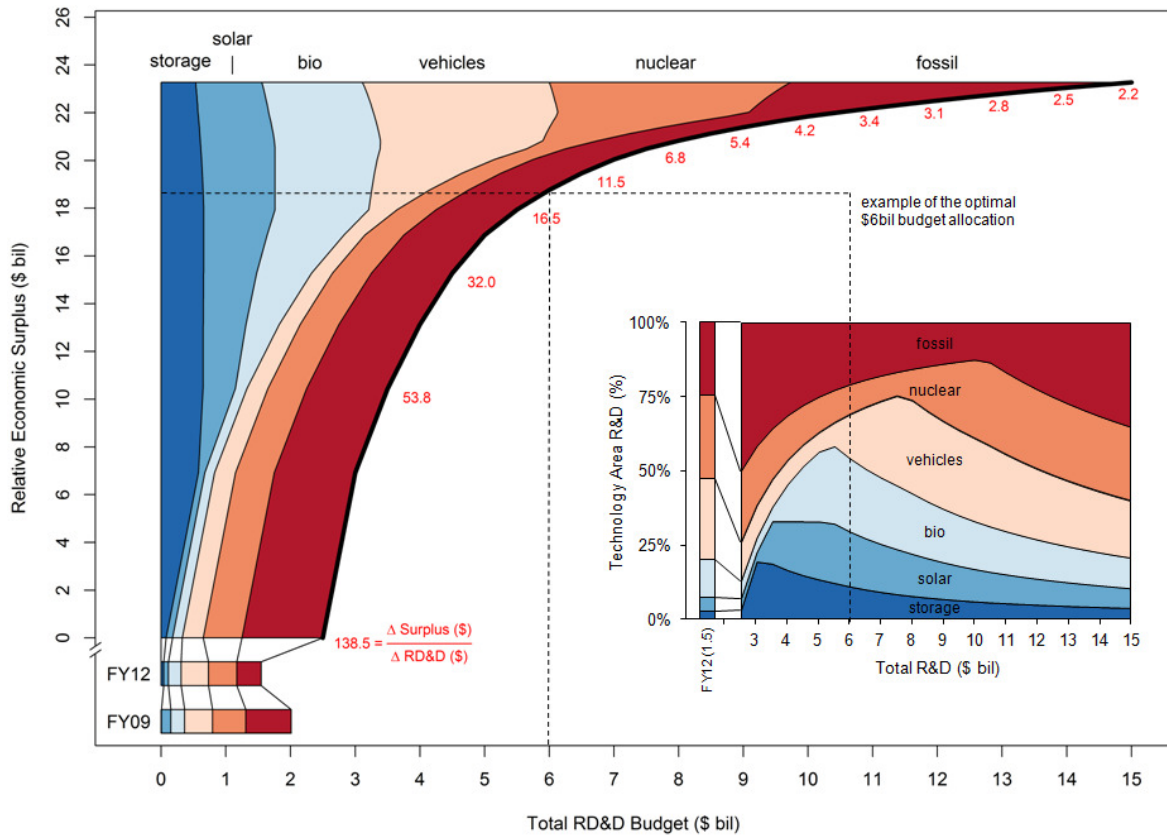
First, there are decreasing marginal returns to R&D. The incremental return to 2030 economic surplus from R&D investments is substantially larger at low levels of R&D than at high levels. In the no policy case, results show that incremental returns to R&D when \$2.5 billion in R&D funding is allocated optimally are \$139 in economic surplus per year in 2030 for an additional dollar of yearly R&D funding allocated to the technology areas with the highest marginal returns at that budget level. Further, there are monotonically decreasing marginal returns to optimally-allocated R&D funds over the full range considered, a subjectively-defined “feasible range” bounded by the estimates of the expert R&D scenarios. However, these results also indicate that there are positive expected returns to economic surplus from R&D even at the highest end of

the range of R&D budget levels considered. This result implies that there are R&D portfolios for the six areas investigated with total budgets greater than \$15 billion (more than 10-times Fiscal Year 2012 levels) that can be justified solely on expected gains to economic surplus realized by 2030. If extrapolated to budgets beyond the range considered, this implies an “optimal” R&D level beyond \$15 billion at which the expected marginal benefits to economic surplus equal the marginal cost of R&D investment. However, this method is not able to capture the marginal cost of raising public funds through revenue-generating policies, such as taxes. The method also does not capture the welfare effects of other externalities associated with energy use, such as pollution and energy resource import dependency

Second, there is a distinct prioritization of R&D investments by technology area. As the R&D budget expands, the optimal allocation shifts first towards energy storage, then to solar energy, then to bioenergy, then to vehicle technologies, and finally to nuclear and fossil energy. Because there are also decreasing marginal returns in the optimally-allocated budget, this result also implies that as the R&D budget expands, the marginal returns to R&D investments in single technology areas are greatest for energy storage, solar energy, and bioenergy. This prioritization of R&D investments appears to be robust across policy scenarios (optimization for minimizing CO<sub>2</sub> emissions reveals strikingly similar results). This is likely due to the low current investment levels in these technologies and the high expected economic returns to R&D in these areas.

Third, there are important differences between the current U.S. public energy R&D funding allocation and the results of this analysis for R&D budgets close to current levels. Comparing the current allocation to the estimated optimum for a \$4 billion budget, fossil energy, energy storage, and solar photovoltaic technologies are

underinvested in. Of these three technologies, this analysis indicates that the technology area that would yield the greatest marginal return to economic surplus, given the current allocation, is energy storage.

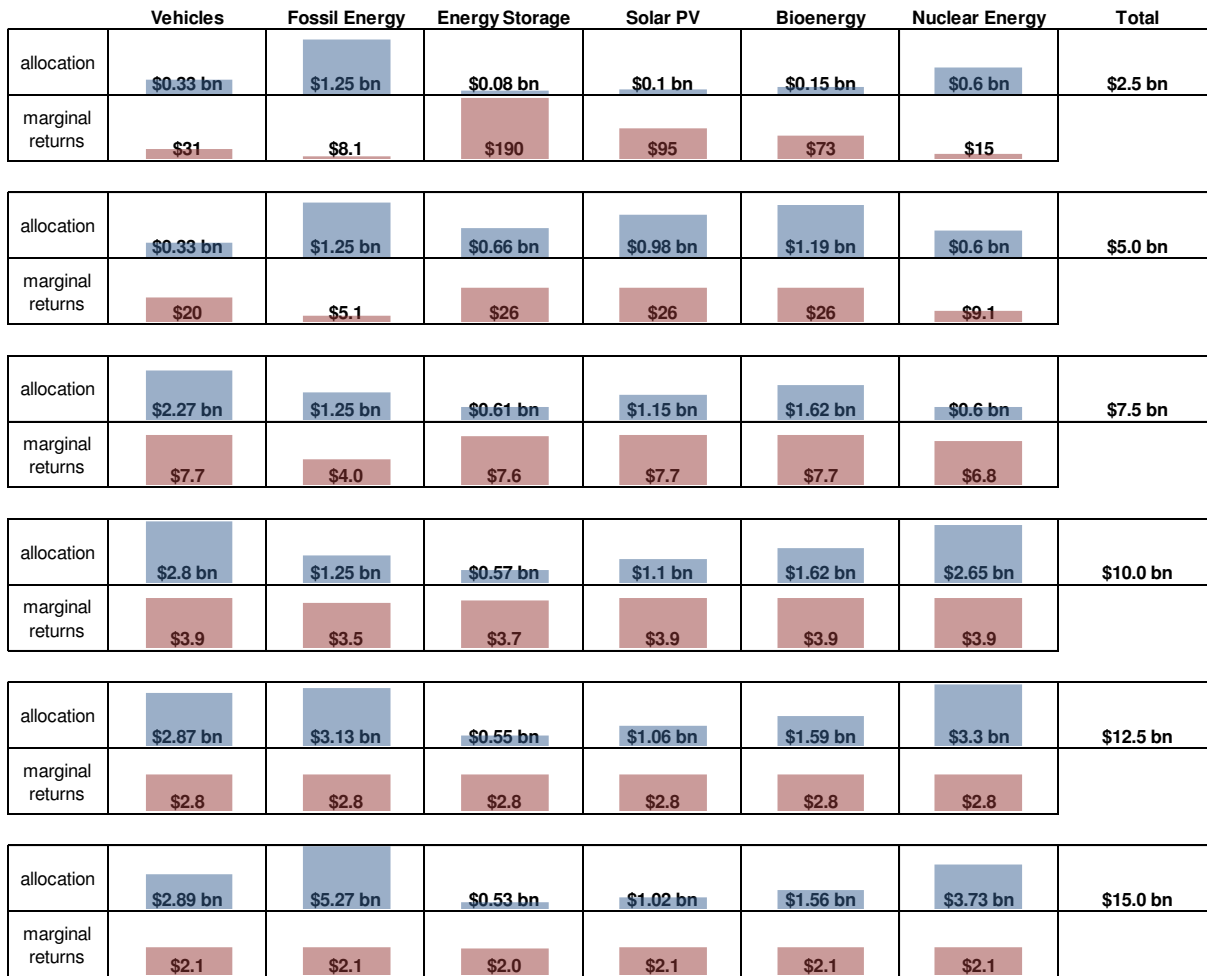


The figure shows the allocation of R&D funding at different R&D budget constraints between \$2.5 billion - \$15 billion per year, relative to the Fiscal Year 2009 R&D budget allocation. The dark black line in the main plot is the maximum expected increase in 2030 economic surplus (above an arbitrary reference point: the expected 2030 surplus in the optimal allocation for the \$2.5 billion budget) that can be attained for a given R&D budget constraint. The red numbers along the black line are estimated marginal returns on investment, calculated by linear approximation to the derivative of the expected 2030 surplus. At the lowest R&D budget considered, \$2.5 billion, the optimal investment mix is 50% fossil, 24% nuclear, 13% vehicles and 3-6% storage, solar, and bioenergy. Because this method is constrained by the range of R&D levels that experts considered, at low levels of total R&D investment estimates are directed towards the technology areas that received expert recommendations for capital-intensive demonstration projects (fossil and nuclear).

**Figure 1.7:** *Optimal R&D Portfolios under a No Climate Policy Scenario*

The range of R&D portfolios considered is tied to the elicitation: to reduce model dependence, expert-elicited relationships between R&D and technology improvements

are not extrapolated, and therefore constrain the analysis to only the R&D ranges that experts explicitly considered. In some cases, this implies that at low and high R&D budgets, optimum portfolios cannot equate marginal benefits across technology areas (see Figure 1.8) because the implied optimum investment level is outside of the range considered by experts.



The figure shows the optimal allocation and marginal returns to economic surplus from R&D at six example budget levels. Corner solutions imposed by the limited range of R&D scenarios considered by experts lead to unequal marginal returns for low budget levels.

**Figure 1.8:** *Optimal Allocation and Marginal Returns by Technology Area*



## 1.4 Conclusions and Discussion

R&D decisions are complex, require integration of multiple assumptions and metrics, and are important for public policy. This chapter shows that current practices in energy R&D decision making in the U.S. could be improved. The new method developed in this chapter is based on a consideration of four design principles for an R&D decision making process. This method also highlights the need for analytic capability that exists between technical experts and funders of R&D portfolios.

The method developed in this chapter advances the current state of methods available in the literature. By operationalizing this method on data from a large set of expert elicitations, the distribution of different types of benefits under different R&D portfolios can be estimated – while fully propagating the elicited uncertainty and accounting for multiple interactions among technologies. Results quantify the aggregate benefits of an R&D portfolio in terms of system benefits (e.g. reduced pollutant emissions and economic surplus), technology benefits (e.g. the deployment of certain desirable technologies), and uncertainty analysis parameters (e.g. statistics of the distribution of other output metrics – such as variance of economic surplus or the probability that an R&D portfolio does not increase CO<sub>2</sub> emissions above certain levels). This method also allows analysts to test the sensitivity of the results to key assumptions, including functional form assumptions, the degree of spillover across technologies, the experts engaged in parameterizing the technology improvement distributions, the R&D investment levels, and other parameters in MARKAL. A transparent presentation of the assumptions that drive the calculation of benefits could help focus new work on important questions that are unresolved.

The method presented in this chapter also improves on current decision making practices at the DOE and several other government organizations, which have demonstrated interest in adopting this type of method. However, there are several possible improvements outside of the scope of this chapter. First, R&D investments could be modeled more realistically. Uncertainties at different innovation stages lead to dynamic time-contingencies in the benefits of R&D investments that occur over time. This chapter models the R&D portfolio decision in a static setting with R&D investment occurring at a single point in time. More recent work has emphasized the dynamic, or process nature of the R&D investments and has analyzed capacity and congestion effects (Terwiesch and Loch, 1999) as well as strategies for search and information gathering (Dahan and Mendelson, 2001; Loch et al., 2001). This implies the need to also consider the sequencing of R&D decisions as part of a dynamic R&D decision making problem (Granot and Zuckerman, 1991; Santen and Anadón, 2014) because aggregate R&D portfolio decisions may have differing impacts depending on the time profile of R&D investments. For example, the large increase and then decrease in federal funding for the National Institutes of Health over the past decade may have been less effective than a slow and continuous increase (Jaffe, 2012), perhaps in part due to the short-run inelastic supply of scientific expertise (Goolsbee, 1998). A dynamic R&D decision-making framework would capture a fundamental aspect of estimating the benefits of an R&D portfolio but would also require additional analytical complexity (on top of a framework that is already quite complex). In particular, the burden on expert participants would likely be very high.

A second direction for future work would be to address the challenge of multiple criteria for assessing the benefits of R&D that may depend on the different values of

stakeholders as well as different technical assumptions. Decision makers may disagree on which criteria to use for assessing the benefits of an R&D program. To be inclusive, decision makers may wish to consider more than one decision making criteria simultaneously (e.g. carbon dioxide emissions, oil imports, and economic growth). There is a long literature in Operations Research on decision making with multiple outcome criteria, sometimes referred to as multi-criteria decision making (MCDM) (Figueira et al., 2005; Greening and Bernow, 2004; Stewart, 1991). There is also literature on directly connecting expert opinions to R&D decision making without the use of an intermediate model of outcomes (Hsu et al., 2003; Liberatore and Titus, 1983). The method presented in this chapter can provide the necessary input data to support MCDM by providing estimates of social benefits measured by several individual criteria.

The approach presented in this chapter could be used in modelling efforts that help justify and improve decision making on R&D beyond the particular application to the DOE that is presented in this work. To generate quantitative estimates of the ex-ante benefits of an R&D portfolio, this method relies on expert judgment, which is inherently subjective in nature. While the inputs to any ex-ante policy assessment require scrutiny, the use of this method to evaluate impacts in a probabilistic fashion can advance public debate in many other cases in which the propagation of uncertainty and the interactions between policy instruments are important. Beyond R&D investment decisions, there are many contexts in which decision making can be supported through the use of a complex model of outcomes with uncertain parametric inputs; for example, earth systems models for decisions about mitigating environmental impacts, structural economic models for decisions about economic incentives, and engineering systems models for decisions about risk mitigation. The approach presented in this chapter has a more generalizable core

that addresses the common challenge in utilizing these types of models of representing uncertainty in parameters and summarizing probability distributions of outcomes that propagate uncertainty (Morgan and Henrion, 1998; Raiffa, 1968).

## 1.5 Acknowledgments

The work in this chapter is part of a broader project on energy technology research and development decision making led by Laura Díaz Anadón. The work presented in this chapter would not have been possible with Professor Anadón's research design, tireless effort, and guidance. This chapter utilizes text from a co-authored manuscript written by myself and Laura Díaz Anadón, but extracts the original work that I designed and conducted myself. However, because this work is part of a larger project, the intellectual contributions of the broader project invariably are also represented here in part. In addition to Laura Díaz Anadón, I would also like to thank the 100 experts who participated in the elicitations for this study, Paul Friley and Savvas Politis of Brookhaven National Laboratory for assistance in running the MARKAL model and Audrey Lee and Melissa Chan for helping design and execute several of the expert elicitations used in this study. I also would like to thank Venky Narayanamurti, Max Henrion, Mort Webster, Valentina Bosetti, Bill Clark, and Joe Aldy for insightful comments and suggestions. This research was supported by the Climate Change Initiative of the Doris Duke Charitable Foundation, the Science, Technology and Public Policy program at the Belfer Center for Science and International Affairs and a grant from BP International Limited entitled "Energy, Climate, and Security Policy." All errors are my own.

## **Chapter 2**

# **The Commercialization of Publicly Funded Science: How Licensing Federal Laboratory Inventions Affects Knowledge Spillovers**

Annual U.S. federal investment in research and development (R&D) has surpassed \$130 billion for the past decade, 40% of which has been allocated to R&D centers funded exclusively by the federal government. For over thirty years national policies have required commercially relevant inventions discovered in federally funded R&D centers to be transferred to the private sector to diffuse knowledge and to promote private sector follow-on innovation. However, there is limited empirical evidence demonstrating the effectiveness of these policies. I quantify the effect of technology transfer on innovation spillovers in the context of patent licensing at the U.S. National Laboratories using data from over 800 licensed patents since 2000. I demonstrate that licensing increases the

annual citation rate to a National Lab patent by 31 – 48%, signifying that licensing is an effective mechanism for inducing innovation. Further, I find that over 75% of follow-on innovation after a patent is licensed occurs outside of the licensing firm, indicating that knowledge spillovers induced by licensing diffuse broadly. These estimates rely on a novel matching algorithm based on statistical classification of the text of patent abstracts. I explore possible mechanisms for the effect of licensing on knowledge diffusion by examining the quality of the patents that cite a licensed patent and rule out the possibility of a strong strategic patenting effect. These results demonstrate that transactions over formal intellectual property to incentivize technology commercialization enhance the benefits of publicly funded R&D by multiplying the creation of knowledge spillovers.

## 2.1 Introduction

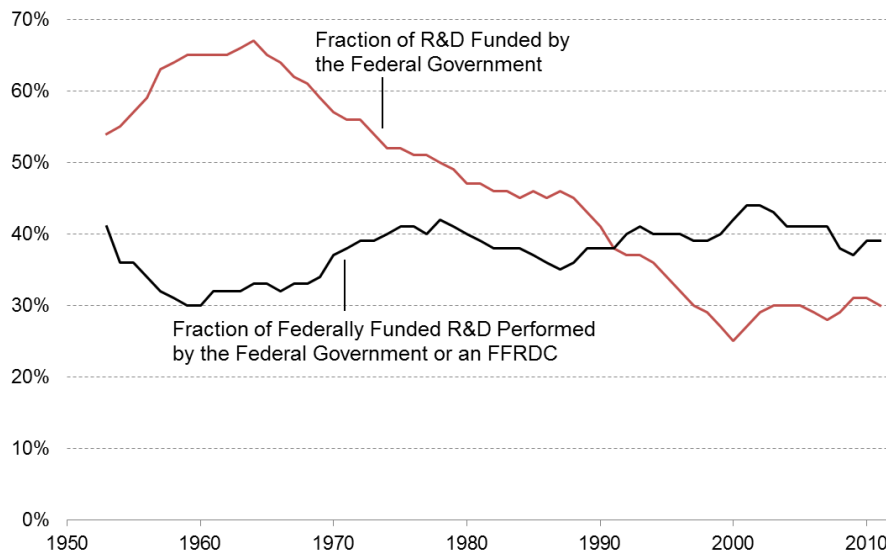
The federal government’s \$126 billion (\$2005) annual investment in research and development (R&D) supports nearly one-third of all R&D in the United States<sup>8</sup> (National Science Board, 2014). R&D is an economically important government function because inappropriable positive spillovers arise when research discoveries serve as the foundation for follow-on innovation (Arrow, 1962; Nelson, 1959)<sup>9</sup>. However, realizing the

---

<sup>8</sup> In 2011, a total of \$424 billion (\$2005) was spent on R&D in all sectors. Of this total, \$126 billion (30%) was provided by the federal government, and of the total federal government expenditure, \$49 billion (39%) was performed directly by the federal government or by an FFRDC. For comparison, total R&D funding spent in universities was \$63 billion in 2011 (\$39 billion of university R&D was provided by the federal government). See Figure 2.1 for the full time series of these ratios from 1953 – 2011. The figure shows that while the share of Federal R&D in total R&D has been declining since the mid 1960’s, the share of intramural and FFRDC R&D in the federal R&D portfolio has been relatively constant. (National Science Board, 2014)

<sup>9</sup> Most simply, R&D can be thought of as the process of knowledge creation. Knowledge is a public good, which implies that markets will undersupply the optimal level of R&D as long as created knowledge induces spillover benefits that cannot be appropriated by the firm conducting R&D. The government also has a role in conducting R&D in fields where it has market power, such as those of strategic national priority, e.g. national defense and space exploration (Cohen and Noll, 1996).

full social value of R&D spillovers, particularly spillovers from publicly sponsored R&D, requires complementary investment in downstream development and commercialization of follow-on innovations (Green and Scotchmer, 1995; Scotchmer, 1991). With this understanding, various policy instruments have been adopted over the past three decades to drive private investment towards commercializing federally funded inventions. These policies allow government sponsored inventors to transfer title of their invention to a private firm for exclusive use, aligning firms' profit incentive with the commercialization of publicly funded technology. Yet because other firms and researchers cannot utilize a technology once it is exclusively transferred, there is a concern that a single firm holding the right to utilize a technology may slow the generation of innovation spillovers (Galasso and Schankerman, 2013; Murray and Stern, 2007; Williams, 2013).



The red line shows the fraction of R&D funded by the federal government from 1953 – 2011. The black line shows the fraction of federally funded R&D performed by the federal government directly or by a federally funded R&D center (FFRDC). Data and definitions adopted from the National Science Board (2014).

**Figure 2.1:** *Time Series of Federal Funding for R&D*

In this chapter, I provide the first large-sample empirical evidence for the relationship between transferring federally funded inventions to the private sector and the subsequent rate of innovation spillovers. Contrary to the concern that exclusively transferred technologies are “held up” within a single firm, I find that technology transfer has a statistically significant and meaningfully large positive effect on the rate of spillovers. This finding is consistent with two socially beneficial effects of technology transfer: (1) information about a technology’s value is revealed when a firm declares its willingness to invest in the technology’s development (Drivas et al., 2014) and (2) knowledge about how to better produce or utilize a technology is created as a firm gains commercialization experience, sometimes referred to as “learning-by-doing.” I further demonstrate that the large majority of spillovers induced by technology transfer diffuse beyond the licensing firm with no evidence that this effect is driven by strategic patenting.

In this chapter I utilize a novel dataset of patents and licensing agreements originating from five U.S. Department of Energy (DOE) FFRDCs. This dataset includes nearly 3,000 utility patents filed from 2000 to 2012, of which over 800 have been licensed. I follow the literature on technological spillovers by using the forward citations to a patent as a measure of induced follow-on innovation (Jaffe et al., 1993, 2000). My empirical strategy utilizes variation in the timing of license agreements with respect to a patent’s age and a novel strategy to match licensed and unlicensed patents. I propose a new procedure for identifying patent matches that utilizes a machine learning algorithm for automated reading of patent abstracts to identify patents of similar technological scope. My proposed approach offers distinct methodological advantages relative to current approaches in the literature that rely on coarse classifications assigned by



patent examiners. I demonstrate the advantages of my approach and show that matching based on classified patent abstracts is less susceptible to omitted variable bias.

Utilizing a difference-in-differences regression design, I estimate that licensing increases the rate at which a patent accumulates citations by 0.22 – 0.34 citations per year above a pre-licensing average of 0.71 citations per year. This amounts to a 31 – 48% increase in follow-on invention induced by licensing. This estimate of licensing-induced innovation is concentrated in a period two to eight years after licensing, as I am not able to accurately measure effects beyond eight years. Alone, this result does not speak to the social benefits of technology transfer since follow-on innovation could be isolated within a licensing firm. Therefore, I examine follow-on inventions stratified by patent assignee and find that in fact, more than 75% of follow-on inventions induced by licensing occur outside of the licensing firm. Taken together, I conclude that licensing a federally funded invention enhances rather than detracts from the realization of the full social value of publicly funded R&D.

Several studies in the literature have examined innovation spillovers at the patent-level (see Section 2.4 for examples); however, only very recently has the literature examined the effect of formal technology transfer arrangements on innovation spillovers. In the most relevant study to this one, Drivas et al. (2014) examine the effect of patent licenses on subsequent citations with a different empirical approach and in a different institutional context, patents managed by the University of California, but find results of the same sign and of marginally greater magnitude to what I find.

Forty percent of federal R&D is conducted directly by the federal government or by federally funded research and development centers (FFRDCs), resulting in over 1,000 newly issued patents per year and a stock of over 4,000 invention license agreements

(National Science Board, 2014). However, despite these totals and multiple policy reforms over the past three decades to enhance technology transfer at FFRDCs, there is limited evidence for the relationship between formal transactions over intellectual property and innovation spillovers from federally sponsored R&D outside of the university context<sup>10</sup>. There are two notable exceptions. Jaffe and Lerner (2001) study the effect of policy reforms and management practices on patenting and technology transfer activities at FFRDCs; however, the authors were limited by data availability and only examined aggregate patterns without specifically studying knowledge diffusion outcomes. Additionally, in a more narrowly focused study, Adams et al. (2003) utilize two surveys on FFRDCs from the late 1990s to conclude that cooperative research and development agreements (CRADAs) between an FFRDC and an industrial lab induce greater patenting by the industrial lab. My contribution in this chapter is to provide the first estimates of the spillover effects of transferring non-university federally funded inventions to the private sector, providing a relevant input into the ongoing policy process of reforming technology transfer policies at FFRDCs<sup>11</sup>.

---

<sup>10</sup> For comparison, 50% of federal R&D is conducted by universities and colleges (National Science Board, 2014), yet patenting in the university context, covered under the Bayh-Dole Act, has received significant attention in the innovation policy literature.

<sup>11</sup> In July 2014, the U.S House of Representatives passed H.R. 5120, the Department of Energy Laboratory Modernization and Technology Transfer Act of 2014, which would reduce certain bureaucratic requirements for technology transfer while giving new authority for DOE FFRDCs to adopt best practice technology transfer activities of other federal agencies. As of October 2014, the bill is in committee in the Senate. (113th Congress, 2014) In a parallel effort, the White House initiated a process in October 2011 under presidential memorandum to accelerate technology transfer at all FFRDCs by requiring federal agencies to develop five-year plans to improve technology transfer goals and metrics, streamline technology transfer processes, and develop regional commercialization partnerships (The White House, 2011). Finally, the President's Fiscal Year 2015 budget request included several relevant new initiatives, described in the President's Management Agenda as priorities for "accelerating and institutionalizing lab-to-market practices," which reflect "the Administration's commitment to accelerating and improving the transfer of the results of Federally-funded research to the commercial marketplace" (The White House, 2014).

Beyond the direct results of this chapter, this chapter also makes two contributions to the broader literature. First, the method that I develop to construct matched control patents based on the text of patent abstracts offers a new tool to reduce omitted variable bias and model dependence in the large literature in innovation economics that relies on matching patents<sup>12</sup>. Because it is too costly for the researcher to read each document, existing studies treat the text of patents as unobservable noise. However, this is clearly an unreasonable assumption made for convenience, as the primary purpose of a patent's text is to disclose novel information. Second, I provide some of the first empirical evidence in the literature on the role of formal intellectual property transactions in enhancing knowledge diffusion in the "market for ideas." While representative data on patent licenses between two private actors is not readily available due to secrecy concerns, I was able to collect scrubbed information for license agreements that involved a public sector partner, making this a unique empirical opportunity to contribute to this literature which has thus far relied heavily on theoretical or sectorally limited evidence (Arora et al., 2004; Hellmann, 2007).

The rest of this chapter is organized as follows: Section 2.2 describes technology transfer at the U.S. National Labs and lessons to be learned from the literature on university technology transfer; Section 2.3 describes the data I use in my analysis; Section 2.4 details the empirical framework of the chapter, including a description of the matching algorithm and text-based classification of patents I develop; Section 2.5 presents the results of my empirical work; and Section 2.6 concludes by putting my results in the context of the broader literature.

---

<sup>12</sup> A few example studies that match patents are Jaffe et al. (1993), Thompson and Fox-Kean (2005), and Singh and Agrawal (2011).

## 2.2 The U.S. National Labs and Technology Transfer

The U.S. National Lab system emerged from the facilities created under the Manhattan Project to build the atomic bomb during World War II. Following the War, the newly created Atomic Energy Commission (AEC) assumed responsibility for the Labs and used its new authority to expand the mission of the Labs to cover fundamental scientific research in nuclear sciences. Following the 1973 Arab Oil Embargo, new legislation and executive actions expanded the mission of the AEC to cover non-nuclear forms of energy R&D, doubled total federal investment in energy R&D, and later shifted institutional management of the Labs from the AEC to the short-lived Energy Research and Development Administration (ERDA). In 1977, the cabinet-level Department of Energy was established to replace ERDA (as well as several other energy-related federal organizations), and assumed responsibility for managing the Labs<sup>13</sup>. Today, the National Lab system includes seventeen labs with a combined \$13 billion R&D budget (FY 2011), 97% of which is provided by the federal government<sup>14</sup>. For comparison, total R&D expenditure at U.S. universities and colleges in 2011 was \$63 billion. (National Science Board, 2014)

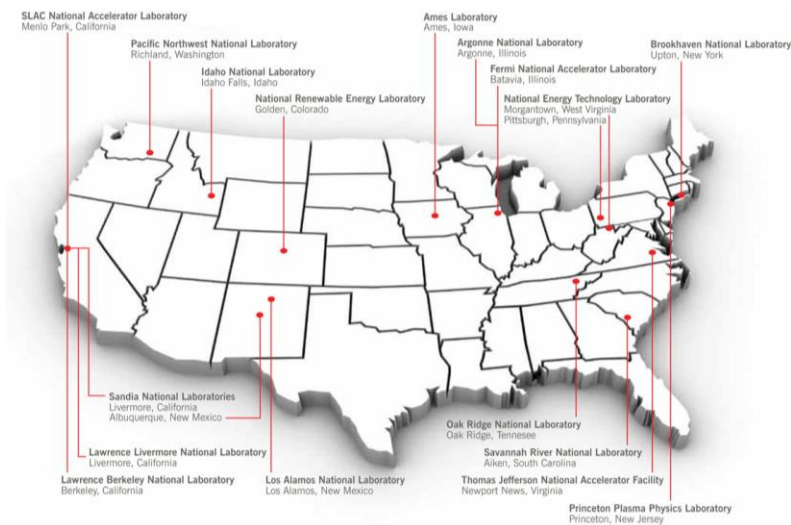
The seventeen National Labs are heterogeneous. First, they are broadly geographically distributed across the country, as shown in Figure 2.2. Some are located near urban centers in close proximity to large research universities while others are in

---

<sup>13</sup> For a general history of the Labs see Westwick (2003) and for a detailed history of each of the seventeen Labs see DOE Office of Scientific and Technical Information (2014a).

<sup>14</sup> The R&D budget of the seventeen DOE labs constitute 75% of all U.S. expenditures at FFRDCs. The remaining 25% of FFRDC R&D is conducted by labs under the responsibility of other Federal agencies or organizations. Prominent examples and their sponsoring agencies are the Jet Propulsion Laboratory (NASA), Lincoln Laboratory (DOD), the National Center for Atmospheric Research (NSF), and the National Cancer Institute (HHS).

remote locations (by design as they were originally secretive nuclear research facilities). Second, the management structure of the Labs varies. While each Lab is owned by the Federal government and must report to DOE, only one of the seventeen Labs, the National Energy Technology Laboratory, is actually operated by DOE. Seven of the Labs are currently operated by a university or university consortium, four are operated by a non-profit R&D company, and the remaining five are operated by an industrial corporation. The Labs also differ in the breadth of their research focus and thus vary in their technology transfer activities. Table 2.1, presents summary statistics for R&D expenditure, patenting activity, and licensing activity for the seventeen Labs disaggregated by their operator type. For a deeper discussion of the implications of different Lab operator types and other Lab management issues, see Jaffe and Lerner (2001), Logar et al. (2014), and Stepp et al. (2013).



DOE is the steward for seventeen National Labs. Within DOE, ten Labs are managed by the Office of Science, three by the National Nuclear Security Administration, and four by the Office of Energy Efficiency & Renewable Energy, the Office of Fossil Energy, the Office of Nuclear Energy, and the Office of Environmental Management. The five labs I study are Brookhaven, Lawrence Berkeley, Pacific Northwest, Sandia, and the National Energy Technology Laboratory. Figure credit: DOE Office of Science (2013).

**Figure 2.2:** Location of the 17 U.S. National Labs under DOE

**Table 2.1: Summary Statistics of R&D, Patenting, and Licensing by Lab Operator**

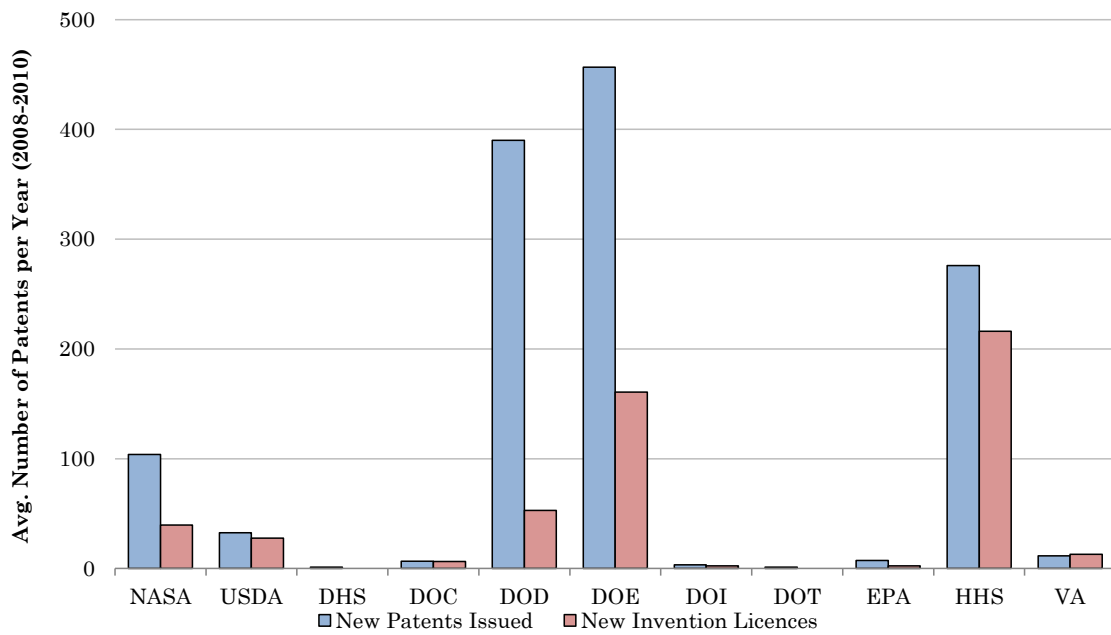
Operator Type	R&D Expenditure per Lab (\$ mil, FY11)	Patents Filed per R&D Expenditure (patents/\$mil)	Patents Licensed per Patent Filed	Patent Licensing Income (\$ mil)	Annual Licensing Income per License (\$/patent)	Licensing Income as Fraction of R&D Expenditure
University (n = 7)	2,453	0.07	11%	13.48	62,235	0.74%
Non-Profit (n = 4)	3,568	0.07	35%	12.80	18,525	0.28%
Government (n = 1)	753	0.01	13%	0.05	4,674	0.00%
Industry (n = 5)	6,569	0.07	13%	12.81	25,305	0.19%
Aggregate (n = 17)	13,343	0.06	18%	39.15	27,484	0.30%
Sample (n = 5)	5,441	0.06	23%	15.68	20,057	0.23%

Summary of R&D expenditure, patenting, and licensing for the National Labs disaggregated by Lab operator type and also showing the sample of five labs examined in this chapter. Summaries of ratios of patenting per million dollars of R&D expenditure and licensing income per license also shown. All variables are three-year averages for 2009-2011 except for R&D expenditure, which is shown at 2011 levels. For comparison, in 2011, all U.S. universities patented at a rate of 0.19 filed patent applications per million dollars of R&D (DOE Office of Energy Efficiency and Renewable Energy, 2012; National Science Board, 2014).

Under the 1980 Stevenson-Wydler Act (P.L. 96-480) and subsequent reforms, all FFRDCs, the National Labs included, are legislatively required to transfer inventions to the private sector<sup>15</sup>. As part of this mission, the FFRDCs have each established technology transfer offices and appropriate a minimum of 0.5% of their R&D budget towards technology transfer, which can include several mechanisms of cooperation with the private sector (e.g. cooperative R&D agreements or “CRADAs”, leasing user facilities—such as bio-refineries and cyclotrons, spin-out company formation, and patent

<sup>15</sup> Important legislation shaping technology transfer policy at FFRDCs include the 1980 Stevenson-Wydler Technology Innovation Act, the 1980 Bayh-Dole Act, the 1986 Federal Technology Transfer Act, the 1989 National Competitiveness Technology Transfer Act, the 2000 Technology Transfer Commercialization Act, and the 2011 America Invents Act. For a discussion of these and other policies shaping the U.S. policy architecture for National Lab and FFRDC technology transfer, see Bozeman (2000), Jaffe and Lerner (2001), Margolis and Kammen (1999), Cannady (2013), and Federal Laboratory Consortium for Technology Transfer (2011).

licensing). In this study I focus on patent licensing because it is both directly measurable in patent-level data and one of the defining activities to understand FFRDC technology transfer, as demonstrated by the central role licensing activity is given in technology transfer performance evaluation. Figure 2.3 compares the patenting and licensing activity enabled by Stevenson-Wydler across each federal agency.



Data is averaged for 2008 – 2010 and is collected from federal agencies by the National Institute of Standards and Technology (2011).

**Figure 2.3:** *Patenting and Licensing of Federally Funded R&D by Sponsoring Agency*

The reforms that followed Stevenson-Wydler slowly changed the way the federal government used intellectual property (IP) to protect government-sponsored inventions. In order to facilitate the transfer of the rights to develop a publicly funded invention effectively, government lab technology transfer offices were given the mandate to quickly and thoroughly apply for IP protection so that eventual licensees could be guaranteed clear rights to utilize an invention in new product development. The long-understood

tradeoff with greater IP protection is that while new inventions are disclosed publicly, access to utilizing new inventions is made exclusive to the right holder. The effect is greater incentive to develop new technologies through the lure of monopoly profits at the societal expense of slowed diffusion of protected technologies, also raising the issue of equitable access to the fruits of innovation. With technology transfer of government inventions to a commercial partner, benefits of publicly sponsored innovations accrue back to the public in the form of access to new technologies and services developed by the commercial partner. Yet, there is a second important channel through which the public benefits from inventions discovered in the organizations it funds: Namely, due to the cumulative nature of innovation (Merton, 1973; Rosenberg, 1982), the introduction of new technologies leads to inspiration for follow-on inventions. Thus, complete evaluation of a policy that affects an innovation system must account for its effect on spurring further inventions (referred to as “spillovers”), a form of positive externality (Scotchmer, 1991). While greater IP protection slows the rate of knowledge diffusion *ceretis paribus*, using IP protection to leverage additional private investment in commercializing technologies that spurs follow-on innovation is a countervailing force.

### **2.2.1 Lessons from University Technology Transfer**

The effect of greater IP protection in the context of university-sponsored research has been thoroughly studied in the context of the 1980 Bayh-Dole Act (Hausman, 2010; Henderson et al., 1998; Mowery et al., 2001, 2002; Wright et al., 2014). Some of the concerns raised in the context of university research also apply to intramural and FFRDC government innovation. Dasgupta and David (1994) summarize one of the most prominent concerns of greater IP protection in these contexts. They argue that promoting greater “industrial transferrability” of basic research findings may induce



short-run benefits through better utilization of existing scientific knowledge but could also have dynamic costs if these activities erode future development of new scientific knowledge. This erosion can occur if researchers are required to divert their effort towards technology transfer activities, for example by dedicating time towards the difficult task of transferring the tacit knowledge that enables the utilization of transferred technology (Arora, 1995; Arora et al., 2004).

An additional dimension of technology transfer of university or government inventions is its effect on the strength of existing IP. Transferring title of an invention to the private sector increases the likelihood that a patent will be litigated for infringement. FFRDCs, like universities, may be unenthusiastic about pursuing costly litigation to enforce their patent rights (Rooksby, 2013), but once title is transferred to a private actor, it becomes in the licensee's interest to pursue litigation for infringement. Finally, there is an ethical concern that inventions discovered in universities or with public funds, once transferred to a single private actor, create benefits inequitably to licensees rather than the general public.

### **2.2.2 National Lab Patent Licenses**

This section briefly describes the general features of National Lab patent licensing agreements. For more detail and a typical sample license agreement see DOE Technology Transfer Working Group (2013). It is difficult to generalize National Lab license agreements, as each license is negotiated individually with particular idiosyncratic terms. Patent license agreements are typically structured to incentivize the licensee to develop the technology (e.g. with performance diligence requirements delineating milestone targets for technology development) while returning a share of profits from commercializing the technology back to the Lab. A National Lab license

agreement typically includes terms for a license issuance fee due when a license is executed, patent cost reimbursement, a minimum annual royalty, and a running royalty equal to a fixed percentage of sales. License agreements can be terminated by the licensee, typically at any point, or by the Lab if diligence requirements or royalty obligations are not met by the licensee. Finally, the U.S. government retains a “march-in” right to re-license an already licensed patent or to use a licensed patent discovered in a National Lab for purposes in the national interest. (DOE Technology Transfer Working Group, 2013; LBNL, Innovation and Partnerships Office, 2014; PNNL, Technology Transfer, 2014)

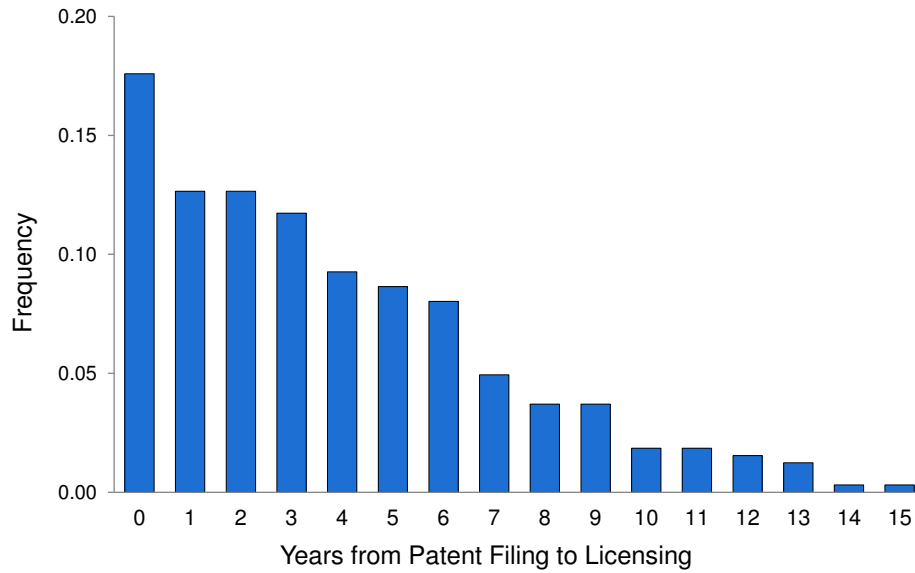
In my dataset, 49% of licensed patents are licensed on an exclusive basis, meaning the Lab agrees to not license the patent to any other interested licensee. The remaining licenses are nearly all partially exclusive, either for a particular field of use<sup>16</sup> or for a certain geographic region. While these non-exclusive licenses could lead to a single patent being licensed multiple times, it is very rare for a single patent to be licensed non-exclusively in such a way that licensees compete for the same market share. One additional distinction between exclusive and non-exclusive licenses is that the right to sublicense a patent to another firm is usually provided for in exclusive license agreements but not non-exclusive licenses. For the empirical section of this chapter, I utilize the first date a patent is licensed to construct the main independent variable.

---

<sup>16</sup> An example of a patent that was non-exclusively licensed for two separate fields of use is patent 6,507,309, “Interrogation of an object for dimensional and topographical information” developed at the Pacific Northwest National Lab. This patent was licensed to a firm to develop millimeter wave body scanners for exclusive application in the fields of aviation, prison, building and border crossing security (these scanners are used extensively in U.S. airports). The same patent was later licensed to another company for a distinct field of use to create body measurements for custom-fit clothing. (Turner, 2004)

In selecting licensees, federal policy requires Labs to give preference to small business licensees and licensees whose production activities are located domestically. In addition, Labs typically must justify their choice of licensee as following from a fair process, although whether or how this is enforced is unclear. In general though, licensing opportunities are advertised on DOE and Lab websites. Although from qualitative interviews, I understand that there is rarely competition among private firms to license the same Lab patent.

Markets for licensing agreements are highly frictional due to the large information asymmetries inherent in transacting over technologies. Despite the existence of patents, which publicly disclose the primary functioning of a technology, nearly all technologies also require additional tacit knowledge possessed by the inventors to be maximally useful (Arora et al., 2004). Because this tacit knowledge is, by definition, not codified and because Lab inventors are typically not aware of the business challenges and technology needs of firms, potential licensees and Lab technology owners face large information asymmetries. This suggests that the role of technology transfer officers is important in finding the suitable matches between available Lab technologies and licensees (Hellmann, 2007). This also helps explain the long tail in the lag between when the Labs file patents and when these patents are eventually licensed (See Figure 2.4).



**Figure 2.4:** *Distribution in the Lag between Patent Filing and Licensing*

## 2.3 Data

In this study, I utilize data on 2,796 utility patents filed between January 1, 2000 and December 31, 2012 and developed at, or in partnership with, five of the seventeen National Labs (Brookhaven National Laboratory, Sandia National Laboratory, Lawrence Berkeley National Laboratory, Pacific Northwest National Laboratory, and the National Energy Technology Laboratory)<sup>17</sup>. These five Labs include at least one Lab in each of the four management structure categories utilized by DOE FFRDCs (government operated, university operated, non-profit operated, and industry operated). Table 2.1 displays summary statistics of R&D expenditure, patenting, and licensing activity for the sample of five Labs I utilize in this study compared to the full set Labs. The five Labs in my sample are representative of the seventeen Labs in terms of the rate of patenting per

---

<sup>17</sup> I contacted the technology transfer offices at the other large National Labs engaged in applied R&D, but was not able to procure sufficient data from these labs to include in my analysis. To account for possible selection issues, I include lab fixed effects in all specifications that do not already include (co-linear) patent fixed effects.

R&D expenditure, but they licensed a slightly higher fraction of their patents (23% instead of 18%) while their average patent license brought in slightly less in terms of royalties (\$20,057 instead of \$27,484). For the empirical section of the chapter, I only utilize data for the five Labs I have patent-level data on licensing.

Data for National Lab patents comes from two DOE databases, the U.S. Energy Innovation Portal, maintained by DOE's Office of Energy Efficiency & Renewable Energy (2014) and DOepatents, maintained by DOE's Office of Scientific and Technical Information (2014b). For each patent, I collect detailed patent-level covariates from two patent databases, the U.S. Patent and Trademark Office's (USPTO) Full-Text and Image Database (U.S. Patent and Trademark Office, 2014) and Google Patents (Google, 2014). These databases allow me to observe the DOE contract number of the R&D agreement the patent was developed under, the name of the inventors and initial assignee, the application, grant dates, priority date, the U.S. and international technology classification, and the text of the patent abstract and claims.

For the five Labs in this study, I obtained comprehensive records of patent license agreements from each Lab's technology transfer office. Of the 2,796 patents in the full dataset, I observe that 877 were licensed between January 1, 2000 and December 31, 2012<sup>18</sup>. For each licensed patent, I observe the date on which the licensing agreement went into effect and whether the license was issued exclusively or non-exclusively.

Table 2.2 presents descriptive statistics for all patents and licensed patents in my dataset aggregated by observations at the patent-level (2,796 patent observations, of

---

<sup>18</sup> Some of the patents in the full dataset may have been licensed prior to January 1, 2000 or after December 31, 2012. However, this is not problematic for my analysis as I limit comparisons to patents filed in similar time windows and the full database covers all unlicensed patents in the time period over which I have licensing data.

which 877 are licensed) and the patent-year-level (27,402 patent-year observations, 9,852 of which are for patents that are licensed between 2000 and 2012).

**Table 2.2:** *Descriptive Statistics for All Patents and Patents that are Ever Licensed*

a) Descriptive statistics at the patent-level

	All Patents					Ever Licensed Patents				
	Obs.	Mean	S.D.	Min	Max	Obs.	Mean	S.D.	Min	Max
Lab = LBNL	2,796	0.16	0.36	0	1	877	0.14	0.34	0	1
Lab = NETL	2,796	0.10	0.30	0	1	877	0.02	0.14	0	1
Lab = PNNL	2,796	0.23	0.42	0	1	877	0.54	0.50	0	1
Lab = SNL	2,796	0.44	0.50	0	1	877	0.27	0.44	0	1
Lab = BNL	2,796	0.07	0.25	0	1	877	0.04	0.19	0	1
Filed Year	2,796	2005.8	3.60	2000.1	2014.0	877	2004.3	3.06	2000.1	2012.8
Grant Year	2,796	2008.9	3.83	2001.3	2014.8	877	2007.6	3.46	2001.4	2014.8
License Year						877	2005.5	3.61	2000.1	2013.4
Grant Delay (yrs.)	2,796	3.09	1.49	0.25	11.17	877	3.26	1.61	0.46	11.17
Total Cites	2,796	7.92	18.51	0	252	877	12.14	22.75	0	236
Ever Licensed	2,796	0.31	0.46	0	1	877	1.00	0.00	0	1
Ever Exclusive Licensed	2,796	0.15	0.36	0	1	877	0.49	0.50	0	1

b) Descriptive statistics at the patent-year-level

	All Patents					Ever Licensed Patents				
	Obs.	Mean	S.D.	Min	Max	Obs.	Mean	S.D.	Min	Max
Year	27,402	2008.9	3.69	2000	2014	9,852	2008.5	3.78	2000	2014
Age from Filing	27,402	5.06	3.69	0	14	9,852	5.53	3.78	0	14
Cites	27,402	0.80	2.32	0	51	9,852	1.08	2.58	0	42
Cites from Unique Assignees	27,402	0.38	1.06	0	33	9,852	0.63	1.50	0	33
Cites from Non-Zero Cited Patents	27,402	0.43	1.62	0	35	9,852	0.59	1.85	0	34
Cites from Above-Median Cited Patents	27,402	0.33	1.37	0	30	9,852	0.45	1.50	0	24
Already Granted	27,402	0.69	0.46	0	1	9,852	0.71	0.45	0	1
Already Licensed	27,402	0.30	0.46	0	1	9,852	0.83	0.37	0	1
Already Exclusive Licensed	27,402	0.18	0.38	0	1	9,852	0.49	0.50	0	1

## 2.4 Estimating a Citation-Based Model of Knowledge Diffusion

In this chapter, I assess the public returns to patent licensing, as measured by the differential citation rates of licensed patents relative to non-licensed patents. To reduce selection bias that may arise from the non-random assignment of licensing status to patents in different technological areas, I carefully match patents based on their filing date, annual citations (pre-licensing), and a novel measurement of their technological scope derived from the text of their abstracts. I then compare citation rates after one of the matched patents is licensed using a difference-in-differences framework.

Measuring knowledge diffusion using technology-level data is difficult due to a lack of available data at a granular level. Patents, however, have proven to be a useful source of data for measuring knowledge diffusion because they include detailed information about the antecedent inventions on which a patent builds. The citations included in a patent also play a legal role by demarcating prior art and thereby limiting the claims of a patent with respect to previous patents, and therefore citations are a noisy but still useful measure of knowledge diffusion. Patent citations are included on the front page of a patent document and are added by inventors, legal counsel, or patent examiners. For a detailed discussion of the role and significance of patent citations in the context of economic research, see Jaffe and Trajtenberg (2002).

The matching method that I propose in this section is novel, but the way in which I estimate and measure spillovers draws heavily on the literature that has examined innovation spillovers at the patent-level. For example, Galasso and Schankerman (2013) examine the effect of patent invalidation on subsequent citations, Singh and Agrawal (2011) ask whether firms develop follow-on innovation through hiring already successful

inventors, and Jaffe et al. (1993) study how geographic proximity between an initial and follow-on inventor affect subsequent citations. As I show below, my proposed matching method has several important advantages for reducing bias and improving the precision of estimates, making it a useful tool to reevaluate some of this earlier literature.

### 2.4.1 Matching

My analysis relies on comparing patents within and across technology application areas. It is well known that the USPTO's patent classification system poorly measures what researchers seek to use the classification as a proxy for (Scherer, 1982). Existing patent classification schemes are not well suited to this task because (1) in the USPTO, patents are not classified by their potential areas of application, but instead are classified by their technical characteristics (Hirabayashi, 2003) (2) the level of granularity in patent classifications is inconsistent across technology areas, and further, within a single class there may be substantial heterogeneity across patents (Thompson and Fox-Kean, 2005), (3) classifications are continuously revised over time, (4) classification relies on idiosyncratic decisions by patent examiners and exploratory analysis reveals that very similar technologies (even pairs of patents that are co-licensed by the same specialized company) are not consistently classified<sup>19</sup>. In addition,

---

<sup>19</sup> One striking example of the idiosyncratic nature of USPTO classifications is shown by the example of the Combustion Controls and Diagnostics Sensors technology developed at the National Energy Technology Laboratory in the early 2000s. This technology involved an initial patent, 6,429,020 titled "Flashback detection sensor for lean premix fuel nozzles," which was filed in June 2000 and granted in August 2002. This patent was followed up with a continuation in part patent, 6,887,069 titled "Real-time combustion controls and diagnostics sensors (CCADS)" which was filed in September 2001 and granted in May 2005. Demonstrating the similarity of these two patents, they shared three inventors, were jointly licensed by Woodward Industrial Controls in December 2001, and were the subject of two CRADA agreements between NETL and Woodward. The abstracts of the two patents are extremely similar: Patent 6,887,069 states it is "an apparatus for the monitoring of the combustion processes within a combustion system," and patent 6,429,020 is described as "a sensor for detecting the flame occurring during a flashback condition in the fuel nozzle of a lean premix combustion system." However, they were given two different primary classifications by the USPTO. The initial patent, 6,429,020, was given the primary classification 436/153. Class 436 is described as "a generic class for ... process[es] which involve a chemical reaction for determining



inventions, particularly high-value breakthrough inventions, very often involve the combination of technologies from distinct fields (Arthur, 2009; Fleming, 2001; Weitzman, 1998), and thus may span multiple USPTO categories. While patents can be placed in multiple categories, they must also have a single declared primary category. In most empirical studies, the single primary category is used for analytic traction. In addition, because of finite sample size, the coarse nature of patent classifications implies that studies that rely on matching have to discard a large number of patents due to the lack of suitable matches. All matching approaches applied to patents will result in a substantial fraction of discarded observations, as by definition, each patent must be “novel,” but coarse measures like the USPTO classification make identifying more similar patents difficult.

Fortunately, patent documents do contain a plethora of information concerning the underlying innovation (that is explicitly what they are designed to do). Patents often contain pages of text and figures, but in previous research, actual reading of the text of patents has proved too time-consuming and too substantively demanding for social science researchers use to determine which patents are most similar in statistical analysis with large sample sizes<sup>20</sup>. Therefore, studies in the innovation literature that utilize patent data have not widely incorporated the fundamental information in patents. This leads to greater omitted variable bias and model dependence as there will

---

qualitatively or quantitatively the presence of a chemical element, compound or complex,” and its 153 subclass specifies that “measurement of electrical or magnetic property or thermal conductivity ... of an ionized gas.” The continuation patent, 6,887,069, was classified not just in a different subclass but also in a different class, 431/12, which is described as “processes of combustion or combustion starting,” and its subclass specifies “processes controlling the supply of fuel or air discharged into the combustion zone.” Admittedly, the two patents do share three classes designated by at least one of them as a secondary class, but neither patent contains the subclass of the other’s primary subclass.

<sup>20</sup> One notable exception is Scherer (1982) who, with a team of four engineering and chemistry students, read and classified 15,000 quasi-randomly selected patents.

be potentially greater endogeneity concerns in these studies due to a lack of meaningful observables on which to account for selection bias.

Because of the issues of using USPTO classifications and recent development in automated content analysis in computer science, I am able to implement a more sophisticated patent classification algorithm. I classify the patents in my dataset using a machine learning algorithm based on the textual content of the patent abstracts. I use the Latent Dirichlet Allocation (LDA) algorithm, which uses a Bayesian model of word co-occurrence, to classify documents into endogenously defined technology topic areas (Blei, 2010; Blei et al., 2003; Blei and Lafferty, 2007).

#### **2.4.1.1 Topic Modeling**

Using text as a primary data source in a causal inference framework is not straight forward. In particular, it is difficult to discern the difference between changes in behavior and changes in the way people use language to *describe* a particular behavior. One advantage of using data from the abstracts of patents, is that innovators themselves tend to write patent abstracts (lawyers are typically more involved in writing the claims in the body of a patent); therefore, it is less likely that the abstracts contain strategically motivated language. In addition, because of the legal function patents play and because of the U.S. common law system (based on precedent), the use of language within patents may be more stable than other corpuses with long time series.

Text-based analysis of natural language has a strong legacy in computer science and statistics<sup>21</sup>. More recently, advances in computational power, the growing acceptance of text-based data in social science research, and the digitization of text sources have led to

---

<sup>21</sup> See, for example, Mostellar and Wallace (1964) for an early text classification analysis in statistics.

a proliferation of text-based analyses in the social sciences (Alexopoulos, 2011; Grimmer and Stewart, 2013; A few recent examples include: Hoberg and Phillips, 2010; Hopkins and King, 2010; Kaplan and Vakili, 2014; King et al., 2013; Quinn et al., 2010). Text-based analysis relevant for social science research has developed several different approaches for classification of documents into similar (predefined or unknown) categories<sup>22</sup>.

In this chapter, I classify patents based on the text of their abstracts to account for otherwise-unobserved heterogeneity in the technological scope of an invention. Text-based classification requires the selection of the appropriate method. Because technologies arise through the recombination of existing technologies (Arthur, 2009; Fleming, 2001; Weitzman, 1998), they can often span the multiple technological categories of their predecessor technologies, making classification methods that allow for multi-category membership well-suited. Further, patents are highly technical documents, which gives multi-membership models greater power in modeling the facets of individual documents (Quinn et al., 2010). Finally, the complexity of patents greatly increases the cost of implementing a supervised method that would require human coding. Therefore, I have chosen the Latent Dirichlet Allocation (LDA) model (Blei, 2010; Blei et al., 2003), which is an unsupervised mixed-membership classification method utilizing a Bayesian machine learning algorithm. A notable disadvantage of unsupervised classification is that the resulting grouping are difficult to interpret since

---

<sup>22</sup> Depending on the complexity of the classification task, text-based analysis can be supervised (meaning human input is required to “teach” an algorithm which documents belong in a category) or unsupervised (meaning the algorithm incorporates a method to “learn” how documents should be divided into categories). Quinn et al. (2010) and Grimmer and Stewart (2013) provide guidance for selecting appropriate classification methodologies in the context of political science research.

they are not pre-defined. However, my main objective in classification is to identify similar patents which doesn't require directly interpreting classes of patents.

Abstractly, the LDA model estimates two simultaneous types of probability distributions. First, for the corpus of documents as a whole, LDA estimates a specified number of latent "topics," representing the likelihood that words co-occur within a document. Each topic is a probability distribution over all words in the vocabulary, with greater probability weight assigned to words that are more likely to occur when that topic appears in a document. LDA requires the number of topics to be pre-specified by the researcher; greater number of topics gives more granular information about the structure of a document. The second set of distributions LDA estimates are the probabilities that any one word in a document originates from one of the topics. For this chapter, the relevant output of the LDA model is a document-level probability vector of topic frequencies. These frequencies provide a continuous measure of a patent's substantive content and can be used to assess the topical similarity of two different patents.

The hierarchical classification of LDA provides useful advantages in a causal inference framework relative to alternative approaches, such as word-frequency clustering (e.g. k-means), that directly classify documents based on word frequencies. First, LDA allows the researcher to control the granularity of the classification by choosing the number of topics, giving the researcher a principled method to determine the specificity of patent classification, and therefore the closeness of matches, in a way that scales with the sample size. The frequency of topics also gives a chosen number of covariates to describe a patent, and thus can help separate meaningful words and combinations of words from meaningless words (whereas direct word frequencies give an

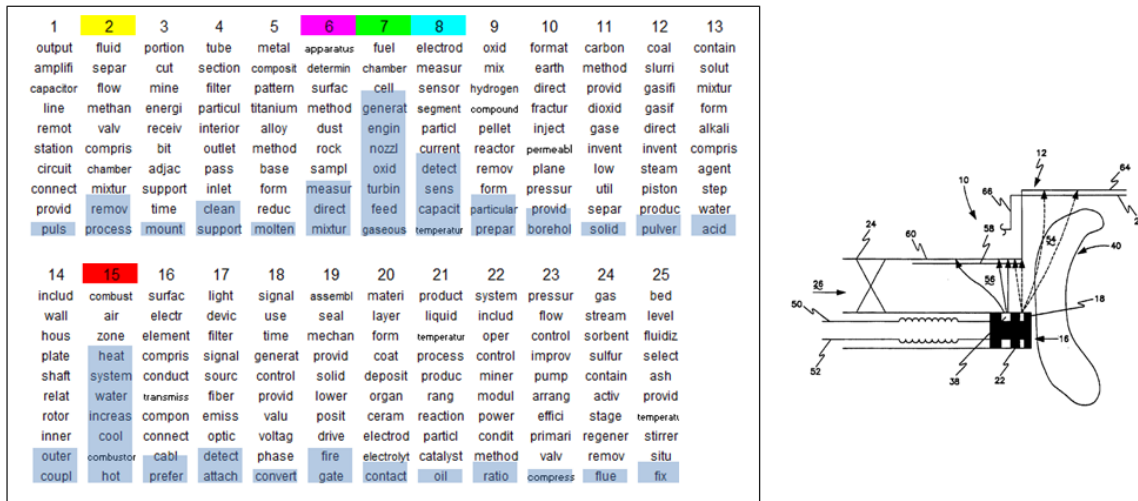
intractable number of covariates, often more than the number of documents). Second, the hierarchical structure provided by the topic distributions allows for words to take on distinct meanings depending on the words they are likely to co-occur with. For example, the word “compound” is likely to have a distinct meaning mean it is used with words concerning chemical compounds compared to when it co-occurs with words that describe inventions with two compounded sub-components.

In Figure 2.5, I provide a visual representation of the LDA model applied to classifying the subset of patents in my dataset from the National Energy Technology Lab. In the example, I implement the LDA model with 25 topics, and show, for a single patent, the document-level distribution of topics, the ranking of most likely words to occur within the topic, and the classification of words within the document to individual topics.

**U.S. Patent 6,887,069 Real-time combustion controls and diagnostics sensors (CCADS)**

The present invention is directed to an apparatus for the monitoring of the combustion process within a combustion system. The apparatus comprises; a combustion system, a means for supplying fuel and an oxidizer, a device for igniting the fuel and oxidizer in order to initiate combustion, and a sensor for determining the current conducted by the combustion process. The combustion system comprises a fuel nozzle and an outer shell attached to the combustion nozzle. The outer shell defines a combustion chamber. Preferably the nozzle is a lean premix fuel nozzle (LPN). Fuel and an oxidizer are provided to the fuel nozzle at separate rates. The fuel and oxidizer are ignited. A sensor positioned within the combustion system comprising at least two electrodes in spaced apart relationship from one another. At least a portion of the combustion process or flame is between the first and second electrodes. A voltage is applied between the first and second electrodes and the magnitude of resulting current between the first and second electrodes is determined.

Filed: Sept. 18, 2001. Granted: May 3, 2005. Patent application licensed exclusively to Woodward Industrial Controls on Dec. 28, 2001.



In this example, I show how U.S. patent 6,887,069 was classified under and LDA model with 25 topics applied to the subgroup of patents in my dataset from the National Energy Technology Lab. The patent abstract is shown at the top, with individual words highlighted in colors corresponding to the five most frequent topics for this document. In the box below the abstract, the topic distributions for the corpus are shown with words within topics arranged in descending order of likelihood to occur within the topic. Overlaid on top of the word-within-topic distributions is the distribution of topics within this particular patent, shown in blue bars. This distribution is the key output of the LDA model for my analysis as it gives a continuous measure of the patent abstract’s substantive content.

**Figure 2.5:** Example Classification of a Patent using the LDA Model

While the LDA model has been applied to classifying patents before (see, for example, the recent work of Kaplan and Vakili (2014) and Venugopalan and Rai (2014)), to my knowledge, this is the first work that incorporates LDA-based classification to account for otherwise-unobserved heterogeneity in a model of causal inference. The LDA model, including full mathematical representation, is described in detail in Appendix B.1.

### 2.4.1.2 Matching Implementation

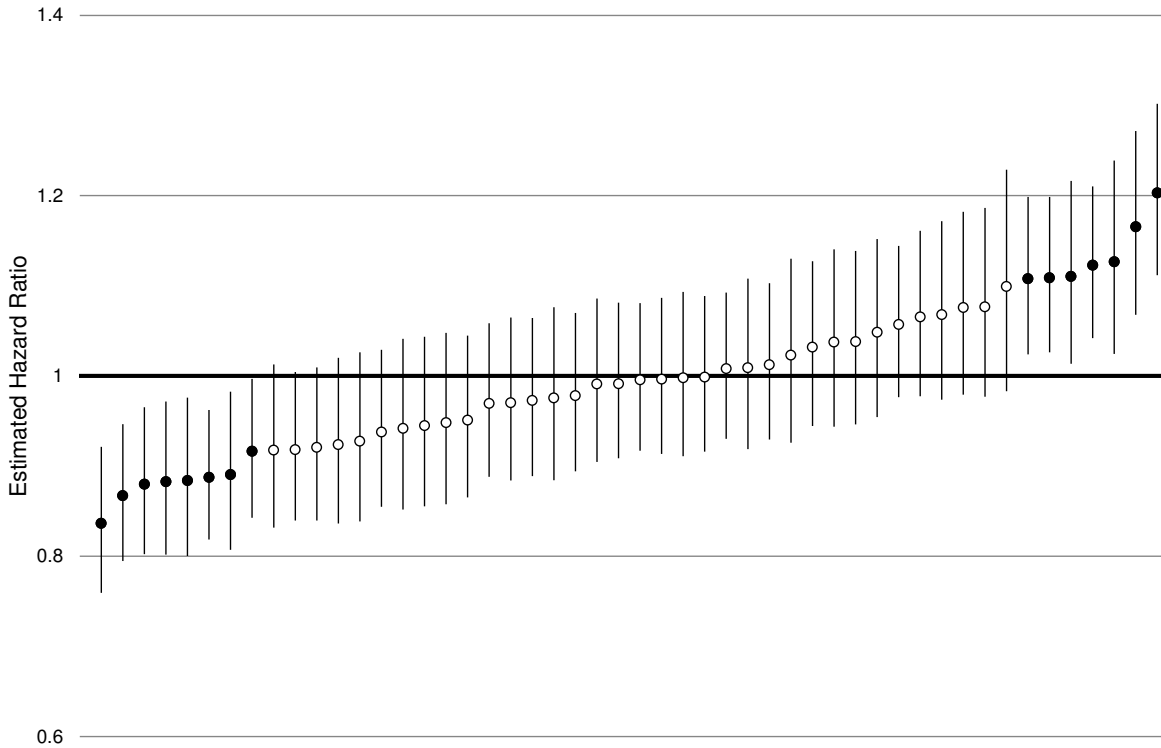
The matching method I implement has several steps which seek to vastly improve the number of observable characteristics extracted from the underlying patents. There are several methodological choices required to implement a matching method with many types of covariates of different relative importance. I describe alternative methods for matching licensed to unlicensed patents and implement several alternative algorithms to assess robustness. These choices apply different combinations of exact matching, coarsened exact matching (CEM) (Iacus et al., 2012) and nearest neighbor matching on available covariates and estimated balancing scores derived from the text-based classification described in Section 2.4.1.1. In all approaches, I apply coarsened exact matching to filing date and average annual pre-licensing citations.

I consider three approaches to matching patents on their estimated topical structure from the LDA analysis: I consider matching patents based on a balancing score that is a function of the estimated topic proportions, direct Mahalanobis distance between estimated topics proportions, and the USPTO classification. For the methods that rely on the LDA classification, I estimate matches with a 25-topic and 50-topic model. The remainder of this section refers to Table 2.8, where regression results from models described in Section 2.4.2 are estimated with different matching approaches.

In the balancing score approach, I match licensed and unlicensed patents based on a balancing score that captures the probability of a patent being licensed in a panel data setting. A filed patent can be licensed at any point and I have information on the timing of each patent's filing and license; therefore, the simple propensity score of being licensed does not fully utilize the information available to estimate licensing probability. To construct a more appropriate balancing score, I implement a Cox proportional hazard

(Cox, 1984) model to estimate the (constant) hazard that a patent is licensed in any one year as a function of a patent's topical structure. The predicted hazard is a balancing score appropriate for panel data where treatment occurs at different points for observations. The use of the predicted hazard as a balancing score is analogous to propensity score matching in cross-section data or panel data where treatment occurs at the same time. (Stange, 2011) I also include lab fixed effects in the hazard regression to account for heterogeneity in lab ability to license. This allows me to utilize the information about the delay between a patent being filed and licensed as a key input for determining licensing probability. Matching on the estimated hazard ratios is assessed in specifications (1) and (2) in Table 2.8. The outcome of the hazard regression is predicted hazard ratios for each topic in the topic model, which can be used to calculate expected hazards of being licensed for each patent. Figure 2.6 displays the range of topic coefficients estimated in the hazard regression which could serve as an input to such a calculation. The results of calculating patent-level predicted hazard based on these coefficients could be useful to technology transfer offices seeking to identify technology areas more promising for licensing. The figure also reveals that the majority of topics are not significant predictors of licensing and thus shouldn't strongly affect how patents are matched – a nuance not captured by Mahalanobis distance matching.





Estimated hazard ratios can be used to predict how likely a given patent is to be licensed based on the classification of its abstract. Topics are arranged from lowest to highest hazard, where a hazard ratio of 1 means no effect on probability of a patent being licensed. 95% confidence intervals on estimated hazard ratios are shown for each topic with statistically significant estimates at the 5% level indicated with filled-in circles. These estimates inform the predicted hazard ratios for each patent, the balancing score used for matching in the preferred specification.

**Figure 2.6:** *Estimated Hazard Ratios for Each Topic*

The Mahalanobis distance matching approach simply takes the distance between the vector of estimated topics and finds nearest neighbor matches. I implement both the Cox proportional hazard model and the Mahalanobis distance matches on the panel of patents using the log of the estimated topic proportions from the topic model as the covariates. I use the log of estimated topic proportions to normalize the distribution of topic frequencies<sup>23</sup>. Matching on the Mahalanobis distance is assessed in specifications (3) and (4) in Table 2.8.

---

<sup>23</sup> The distribution of topic frequencies is strongly right-skewed, which is a desirable property of the LDA model as it demonstrates stronger differentiation of topics.

To reduce the noise from small document-level estimated topic proportions, I also apply a calipers approach that sets all estimated topic proportions below the 90<sup>th</sup> percentile of topic proportions to zero. These approaches utilize only the estimated topic proportions that describe a substantial proportion of the patents. I then perform CEM on a binary indicator for the patent having an above-90<sup>th</sup> percentile estimated proportion of each topic and then resolve CEM matches with nearest neighbor matching based on the estimated hazard (specifications (5) and (6) in Table 2.8) and with Mahalanobis distance (specifications (7) and (8) in Table 2.8).

In the USPTO classification approach, I match patents exactly on their primary assigned class, ignoring secondary classes and subclasses to increase the number of matches. Still, while matching for topical scope using the hazard ratio or distance metrics preserves more than 9,000 patent-year observations, matching just on the primary USPTO classification reduces the number of observations to below 3,000 patent-year observations due to the lack of suitable controls. This again highlights the strength of the LDA classification approach in allowing for the granularity of topic matches to be defined relative to the sample size. Within PTO classes, I choose one-to-one matches randomly (specification (9) in Table 2.8) and also according to nearest neighbor matches in terms of the estimated hazard (specifications (10) and (11) in Table 2.8) and Mahalanobis distance (specifications (12) and (13) in Table 2.8).

In what I call the “preferred matched sample” I apply CEM matching on patent filing year and pre-license average annual citations. I then use the 50-topic LDA model and match patents one-to-one based the topic-dependent estimated hazard of being licensed.

A simple assessment of the matches is shown in Table 2.3, which gives naïve difference-in-difference estimates based on conditional means, using the matching to

control for pre-treatment heterogeneity in observables. Assessment of balance after matching under the preferred specification is shown in Table 2.4, which shows that filing year, grant year, priority year, and grant delay are all statistically distinct for licensed and unlicensed patents in the full sample, but balanced in the preferred matched sample.

**Table 2.3:** *Conditional Means of Annual Citations Approximating Diff-in-Diff*

- a) Citations per year for patents that are never licensed and patents that are licensed during the time period of the panel

	Pre-Licensing	Post-Licensing	Post - Pre
Never Licensed	0.85	n/a	n/a
Ever Licensed	1.40	1.59	0.20
Licensed - Unlicensed	0.54	n/a	n/a

- b) Citations per year for patents in the preferred matched sample before and after licensing (or licensing of matched patent). The difference-in-difference estimate in the bottom right cell approximates the regression results shown below but without additional controls.

	Pre-Licensing	Post-Licensing	Post - Pre
Never Licensed	0.70	0.83	0.14
Ever Licensed	0.71	1.53	0.82
Licensed - Unlicensed	0.01	0.70	0.69

**Table 2.4: Balance Check**

	Licensed Patents	All Unlicensed Patents	Difference (Full Sample)	p-value
Filing Year	2003.766 (0.103)	2005.855 (0.083)	2.088 *** (0.141)	<0.0001
Grant Year	2007.022 (0.117)	2008.865 (0.088)	1.843 *** (0.152)	<0.0001
Priority Year	2002.114 (0.106)	2004.684 (0.097)	2.570 *** (0.161)	<0.0001
Grant Delay (Days)	1190.032 (19.864)	1099.118 (11.843)	-90.914 *** (22.073)	<0.0001
Observations	877	1,919		

	Matched Licensed Patents	Matched Unlicensed Patents	Difference (Matched Sample)	p-value
Filing Year	2003.365 (0.139)	2003.474 (0.142)	0.109 (0.199)	0.5850
Grant Year	2006.756 (0.173)	2006.739 (0.172)	-0.016 (0.244)	0.9473
Priority Year	2002.422 (0.156)	2002.519 (0.155)	0.096 (0.220)	0.6607
Grant Delay (Days)	1231.788 (9.026)	1188.940 (28.679)	-42.847 (40.806)	0.2940
Observations	404	404		

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*,  $p < 0.001$

Comparison of conditional means of key patent-level variables for licensed and unlicensed patents in the full sample and the preferred matched sample.

### 2.4.1.3 Evaluating Matching Covariates with Rosenbaum Bounds

Unbiasedness of matching estimators for causal inference depends on the conditional independence condition, which requires that after conditioning on the observable covariates used in matching, treatment assignment is independent of potential outcomes, or “as good as randomly assigned.” The fundamental identification concern with matching is therefore whether there are unobserved factors that affect treatment

assignment and bias causal estimates. An unobserved factor will bias a causal estimate only if it is correlated with the outcome variable and with treatment assignment. With this motivation, Rosenbaum (Rosenbaum, 2005, 2002), proposes assessing sensitivity to the conditional independence assumption by calculating the maximum value of the quantity  $\Gamma$  that would make a causal estimate of interest no longer statistically significant, where  $\Gamma$  is the largest difference in odds of treatment for two units with the same value of observed covariates. In effect,  $\Gamma$  can be thought of as a measure of the maximum covariation in unobservable covariates and treatment assignment that preserves the conclusion of a hypothesis test based on a matching estimator. By assessing sensitivity of a matching estimator to  $\Gamma$ , overall sensitivity to all unobservables can be assessed, thus providing a useful assessment of the conditional independence assumption.

I compare several matching approaches using Rosenbaum Bounds to assess sensitivity of matching to violations of the conditional independence assumption. I do this on a subset of my dataset that provides a more straightforward comparison based on binary treatment and a single dependent variable. The objective of this analysis is to compare matching approaches that control for the technological scope of patents to account for non-random “assignment” to being licensed. For this section, I only examine licensed patents that were licensed in the same year that they were filed. I further restrict this subsample to patents with PTO classifications in which there is at least one licensed and one unlicensed patent (for fair comparison across methods). The final size of this subsample contains 140 licensed patents, for which I attempt to find a match in a pool of 1,131 unlicensed “control” patents. I search for non-licensed patent matches using seven approaches and calculate the average change in citations for licensed patents due to licensing, which can be thought of as an average treatment effect on the treated

(ATT). The approaches control for technological scope of patents in different ways, and thus are biased to different degrees depending on how well they capture the unobservable factors driving whether a patent is licensed. The key question I want to understand with this analysis is whether capturing the technological scope of a patent using the topic modelling approach is less susceptible to omitted variable bias relative to approaches that use the examiner-assigned primary PTO classification (as previous work in the literature has used).

For each of the seven approaches, I present the average treatment effect (on the treated) and its standard error, and the Rosenbaum Bound that would make the average treatment effect not statistically different from zero. It should be noted that the estimates presented in this section are not meant to be directly compared to the estimates in the remainder of the chapter as these estimates do not take advantage of the panel nature of the data (and therefore do not include important control strategies that I utilize elsewhere, such as patent fixed effects).

Table 2.5 summarizes the results from the analysis of Rosenbaum Bounds. Without matching, this subsample of the data contains licensed patents that receive 0.71 (S.E. = 0.11) additional citations per year compared to unlicensed patents (which receive 0.46 citations per year). In the simplest approach, I match licensed and unlicensed patents based on a propensity score model with two sets of fixed effects for the patent's filing year and its originating Lab. With just this simple matching, the estimated difference in citations is 0.24 (S.E. = 0.36). However, this effect is not statistically significant (so there is no relevant Rosenbaum Bound). Next, I add onto this propensity score model by including fixed effects for each primary PTO classification. This increases the treatment effect estimate to 0.62 (S.E. = 0.16). The associated Rosenbaum Effect for this matching

procedure is 1.34, which can be interpreted as follows: if two patents have the same probability of being licensed based only on their filing year, originating Lab, and PTO classification, unobservable factors can cause at most a 1.34 factor difference in the odds of being licensed before the estimated difference in citations is no longer statistically different from zero at the 5% level (for a two-sided test).

**Table 2.5: Sensitivity Analysis with Different Matching Approaches**

Matching Model	ATT Difference in Citations (Licensed - Unlicensed)	Std. Error	t- Statistic	$\Gamma$
No Matching	0.71	0.11	6.66	n/a
Baseline*	0.24	0.36	0.67	n/a
Baseline + PTO Class	0.62	0.16	3.97	1.34
Baseline + 25 topic model	0.58	0.18	3.21	1.64
Baseline + 50 topic model	0.63	0.17	3.82	1.45
Baseline + PTO Class + 25 topic model	0.55	0.18	3.11	1.20
Baseline + PTO Class + 50 topic model	0.69	0.20	3.43	1.59

\*Note: Baseline propensity score model includes filing year and lab fixed effects. Standard errors and t-statistics are heteroskedasticity-consistent as proposed by Abadie and Imbens (2006) using 20 nearest neighbors.

Average treatment effects on the treated units (ATT), standard errors, t-statistics, and Rosenbaum Bounds sensitivity parameters for seven matching approaches. Estimates are for a subsample of the data of treated patents that are licensed in the same year that they are filed in (making all citations in the post-licensing period for ease of comparison).

I compare this result to two approaches that use the topic modeling approach described earlier in this section: once using a 25 topic model and once using a 50 topic

model. I include the log of the estimated patent-level topic proportions<sup>24</sup> along with the same three baseline fixed effects in the propensity score model. I find similar average treatment effects with these two approaches, 0.58 (S.E. = 0.18) and 0.63 (S.E. 0.17), respectively. Most importantly, I find larger values of  $\Gamma$  in the analysis of Rosenbaum Bounds: 1.64 and 1.45, respectively. A direct comparison of values of  $\Gamma$  is possible because the effect estimates from the topic modelling matches are either smaller or approximately the same as the PTO propensity score matches. In both cases, the topic models perform better than the PTO matching, and the 25-topic model performs better than the 50-topic model.

Finally, I implement two additional matching approaches that combine the PTO classification with the two topic model outputs. In these models I find treatment effect estimates of 0.55 (S.E. = 0.18) and 0.69 (S.E. = 0.20) with corresponding values of  $\Gamma$  1.20 and 1.59.

The results from this sensitivity analysis highlight several considerations for matching design. First, matching with the topic model with 25 and 50 topics performs better than with the PTO classification. The matching procedure with 25 topics appears to be the most robust to confounders. Second, the 25 topic model performs better than the 50 topic model, but when combined with the PTO classification, this is reversed. This finding exemplifies the bias-variance tradeoff in matching (Black and Smith, 2004). Third, while the topic model performs better than the PTO classification, it is not clear whether combining the two approaches together is preferable, again because of the bias-variance tradeoff. Finally, it is important to note that the matching procedure relying on

---

<sup>24</sup> Taking the log regularizes the topic proportions and makes the distribution of topic proportions across documents close to Normal.



the topic models has a potentially large advantage over the PTO classification matching due to the condition of common support. Matching on the PTO classification requires dropping all patents in PTO classes in which all patents in the sample were licensed or all were unlicensed (in these cases the PTO class perfectly predicts licensing). This could be a potentially large loss of data (in this example, I dropped eight licensed patents out of 148 that met the other criteria for the subsample, but I was fortunate to have a large control group to pull from). The topic modeling matching approach allows the “distance” between any two patents of different technological scope to be compared, and therefore significantly reduces matching issues that arise from the common support condition.

#### **2.4.1.4 Remaining Selection Bias Concerns**

The fundamental challenge in estimating the causal effect of licensing is that patents are selected for licensing by interested agents. This can be problematic for causal identification if patents that are licensed are also cited more, for example because they are of greater value (see Table 2.3, which shows that the conditional mean of annual citations to licensed patents before licensing is 1.40 but 0.85 for unlicensed patents). A more nuanced version of selection bias can occur if patents that are available for licensing are actually licensed at a time when they would be of greater value, for example when complementary discoveries are made elsewhere in the economy. This second form of selection bias shows that controlling for pre-licensing citations, a proxy for patent value, alone may not solve the selection issue if there are secular trends in technology-area value. A third and related issue is that of simultaneity, patent licensing may occur in anticipation of a technology becoming more useful. In this case, patents may appear to be cited more after licensing but they would have had greater citations even without licensing.

Matching reduces pre-treatment imbalance, importantly bias from secular technology trends. However, all matching approaches rely on an identification assumption that all relevant variables determining selection (in this case, being licensed) are observable and controlled for to mitigate selection bias (Heckman and Navarro-Lozano, 2004; Imbens, 2004). Without random assignment of patents to licensing status, it is impossible to rule out selection bias concerns. Nevertheless, the matching approaches I present in this chapter make an important contribution by providing a method to incorporate a large number of relevant characteristics of patents described by the text of the patent documents, previously treated as unobservable. Introducing these additional covariates reduces bias relative to other approaches that ignore the text of the patents as long as the text of the patent abstracts contain information relevant to whether a patent is licensed that is not incorporated elsewhere. This almost certainly has to be the case in this context if we believe licensees carefully select the patents that they license, so my approach reduces selection bias concerns relative to simpler approaches to matching on observables.

In addition to matching, I utilize patent-level fixed effects in my preferred specifications which focuses the identifying assumption on the timing of licensing relative to unobservable factors that differentially affect a patent over time. Patent fixed effects control for all unobservable time-invariant characteristics of patents, such as the patent's inherent value, so the most concerning omitted variable bias must stem from factors that simultaneously affect the timing of licensing and cause citations to licensed patents to relatively increase. Examples include a shock to a patent's value that isn't picked up by the topic model that affects license probability and citation rate, such as the discovery of a complementary technology. I feel confident that this is not likely to be a

strong factor as there are huge frictions in the market to license these technologies. From qualitative interviews with technology transfer offices, Lab employees describe the huge effort and time lags required to market technologies to the right companies, implying that licensing is very slow to respond to secular technology-sector trends.

### **2.4.2 Regression Framework**

I use patent citations to infer the effect that a patent has had on knowledge diffusion; when a patent is cited, I interpret this as the patent in question having spurred a subsequent invention (Jaffe et al., 1993; Jaffe and Trajtenberg, 1999, 1996). In the baseline specifications, the key dependent variable of interest is the annual citations that a patent receives. In each specification, I estimate the causal effect of a patent being licensed on the rate of citations it receives. In this section, I present a set of regression models that build up from simple cross-section and time-series regressions to full difference-in-difference regressions. The specifications that I implement are inspired by previous work that has estimated the causal effect of events on patent or paper-level citation rates (e.g. inventors changing firms, filed patents being eventually granted, and contested patents being ruled invalid) (Drivas et al., 2014; Furman and Stern, 2011; Galasso and Schankerman, 2013; Murray and Stern, 2007; Singh and Agrawal, 2011). In the specifications I discuss below, I also apply the matching approaches described in Section 2.4.1 to preprocess the data to reduce selection bias and model dependence.

The simplest models to quantify the effect of a patent being licensed takes either 1) a cross-section of licensed and unlicensed patents and compares citation rates or 2) a time series of licensed patents and compares citation rates before and after licensing. These two models suffer from selection bias as they do not account for the differential quality of

licensed versus unlicensed patents nor possible secular trends driving both licensing and citations. Nevertheless, I present these models for comparison.

The cross-section regression is estimated with the equation:

$$CITES_{i,t} = f(\psi_L EVER\_LICENSED_i + \delta_{i,t} + \gamma_t + \varepsilon_{i,t}). \quad (2.1)$$

In this equation,  $CITES_{i,t}$  are citations received by patent  $i$  in year  $t$ ,  $EVER\_LICENSED_i$  is a dummy variable equal to 1 if patent  $i$  is licensed and 0 if the patent is not licensed,  $\delta_{i,t}$  are fixed effects for a patent's age (based on the filing year of the patent), and  $\gamma_t$  are fixed effects for the citing year.  $\psi_L$  is the coefficient of interest as it estimates the difference in citation rate for licensed patents relative to unlicensed patents. This regression is estimated using data on the post-licensing period only, using the post-period for unlicensed patents defined by their matched pair that was licensed.

The time series regression is estimated with the equation

$$CITES_{i,t} = f(\psi_{LP} LICENSED_{i,t} + \delta_{i,t} + \gamma_t + \varepsilon_{i,t}). \quad (2.2)$$

$LICENSED_{i,t}$  is a dummy variable that takes the value 1 if patent  $i$  in year  $t$  has been licensed and is 0 otherwise. This equation is estimated only for patents that are eventually licensed.  $\psi_{LP}$  is the coefficient of interest as it represents the change in citation rate for a patent that is licensed relative to its citation rate before it is licensed.

The cross-section and time series regressions are biased due to selection. I estimate these models for a sample of patents where licensed patents are matched to unlicensed patents based on their pre-licensing characteristics. This helps to partially correct for bias by balancing observed omitted variables. However, bias still clearly remains due to unobserved patent-level factors.

A difference-in-difference approach more rigorously estimates causal effects by accounting for systematic time-invariant differences between licensed and unlicensed

patents and patent-invariant differences between patents of different ages. The basic difference-in-difference regression is estimated with the equation

$$CITES_{i,t} = f(\psi_L EVER\_LICENSED_i + \psi_{LP} LICENSED_{i,t} + \psi_P POST_{i,t} + \delta_{i,t} + \gamma_t + \varepsilon_{i,t}). \quad (2.3)$$

$POST_{i,t}$  is a dummy variable equal to 1 if a licensed patent has been licensed by year  $t$  or if the matched licensed patent for a never licensed patent has been licensed by year  $t$ . This approach makes a strong assumption that matching finds comparable pairs, as contrasted with the more lenient assumption that matching achieves covariate balance overall between licensed and unlicensed patents (Rubin, 2006). A natural extension to this model is to add,  $\alpha_i$  matched-pair fixed effects to account for the fixed characteristics of matched pairs:

$$CITES_{i,t} = f(\psi_L EVER\_LICENSED_i + \psi_{LP} LICENSED_{i,t} + \psi_P POST_{i,t} + \alpha_i + \delta_{i,t} + \gamma_t + \varepsilon_{i,t}). \quad (2.4)$$

A more rigorous approach to estimating a difference-in-difference regression relaxes the assumption of strict matched-pair design, instead relying on matching for covariate balance. In this approach, patent-level fixed effects account for all fixed unobserved heterogeneity of individual patents. This regression is estimated with the equation:

$$CITES_{i,t} = f(\psi_{LP} LICENSED_{i,t} + \sigma_i + \delta_{i,t} + \gamma_t + \varepsilon_{i,t}) \quad (2.5)$$

where  $\sigma_i$  represent patent-level fixed effects. As before,  $\psi_{LP}$  is the coefficient of interest. However, in this equation, the coefficient represents the difference in citations from licensing relative to the change in citations for unlicensed patents at similar relative ages.

Finally, the difference-in-difference regression can be disaggregated to estimate yearly effects of licensing, meaning the effect of licensing on citations in specific years prior to and following licensing. This regression is estimated with the equation:

$$CITES_{i,t} = f \left( \sum_{j=1 \dots 10} \psi_{PRE_j} PRE\_LICENSE(j)_{i,t} + \sum_{k=1 \dots 10} \psi_{POST_k} POST\_LICENSE(k)_{i,t} + \sigma_i + \delta_{i,t} + \gamma_t + \varepsilon_{i,t} \right) \quad (2.6)$$

which estimates ten coefficients on yearly difference-in-difference effects prior to licensing,  $\psi_{PRE_j}$ , and ten yearly coefficients after licensing,  $\psi_{POST_k}$ . Matching to improve balance in pre-licensing covariates should effectively make each estimated  $\psi_{PRE_j}$  close to zero, as these coefficients represent differences in citation rates prior to licensing.

Each of these regression models uses annual forward citations as the dependent variable. Most simply and easiest to interpret is a linear model which is well suited to applications with many fixed effects (the incidental parameter problem with non-linear models is not a concern with OLS). However, because citations are a right-skewed count variable (Scherer and Harhoff, 2000), there are several options available to specify alternative functional forms. Non-linear models for count data, such as the negative binomial model, explicitly account for features of count data and can be more appropriate in some contexts. The negative binomial model is a more-flexible extension of the Poisson regression but cannot account for patents that never receive any citations (see Angrist and Pischke (2009) for a discussion of tradeoffs between linear models and non-linear models suited for different types of dependent variables). Finally, for comparability to other studies in the literature, I also estimate a log-linear specification. To avoid the problem of the negative binomial model in accounting for never-cited patents, other studies in the patent literature transform the dependent variable by adding 1 before taking the log (Murray and Stern, 2007). For comparability to the previous literature, I present this functional form, but it is known to be problematic.

## 2.5 Results

This section presents the results from the application of the text analysis, matching, and regression frameworks presented in Section 2.4. First, I present the main regression results from applying Equations (2.1) – (2.6) to the preferred matched dataset of patents. I then examine the extent to which diffusion may be localized to one firm repeatedly innovating. Turning to mechanisms, one concern is that licensing could drive strategic patenting in technology areas that competitor firms now see as more desirable to enter. I account for this by looking at the citation rates of the citing patents themselves, noting that strategic patents are not cited often.

### 2.5.1 Diffusion after Licensing

The results of applying Equations (2.1) – (2.6) to the preferred matched dataset are presented in Table 2.6. Models (1) – (6) in Table 2.6 correspond to Equations (2.1) – (2.6). Model (1) shows that in the post-licensing period, licensed patents receive 0.344 (SE = 0.100) more citations than unlicensed patents. In this model and in models (2) – (4), the post-period for unlicensed patents is defined by the matched licensed patent. A cross-section regression of this form is typically biased because pre-treatment citation rates between licensed and unlicensed patents could differ. However, this is mitigated by the matching algorithm which used pre-treatment citations as a balancing covariate. Nevertheless, this regression may still suffer from other forms of omitted variable bias due to differences in licensed and unlicensed patents.

Model (2) in Table 2.6 shows that licensed patents receive 0.520 (SE = 0.155) additional citations per year after being licensed relative to their pre-license citation rate. Again this regression is biased if there are other factors that occur simultaneously

at the time of licensing. Age and citing-year fixed effects do help reduce some of these concerns, but the lack of a suitable control group in this regression may bias this estimate.

Models (3) and (4) in Table 2.6 present difference-in-difference regressions that now control for heterogeneity in time-invariant differences between licensed and unlicensed patents as well as differences between pre-license and post-license periods. Model (4) adds fixed effects for each matched pair, which controls for a more specific layer of time-invariant heterogeneity. Because the matching algorithm already balanced licensed and unlicensed patents on pre-treatment citations, it is not surprising that the coefficient estimates in these regressions are similar to the results of Model (1) and that the coefficient on `EVER_LICENSED` is close to zero. In model (3), I estimate an effect of licensing of 0.328 (S.E. = 0.091) citations per year without matched-pair fixed effects and in model (4), I estimate an effect of licensing, which includes matched-pair fixed effects, of 0.335 (S.E. = 0.091).

Model (5) extends Model (4) by replaced matched-pair fixed effects with patent-level fixed effects. This relaxes the assumption of matching by allowing for separate estimates at the patent-level instead of the pair-level. As a result, `EVER_LICENSED` and `POST` drop and the coefficient on `LICENSED` remains the difference-in-difference coefficient of interest. I estimate that licensed patents receive 0.223 (SE = 0.066) additional citations after being licensed relative to unlicensed patents over similar time periods. This estimate represents a 31% increase in the citation rate for licensed patents before they are licensed (from 0.71 cites/year for eventually licensed patents before they are licensed).



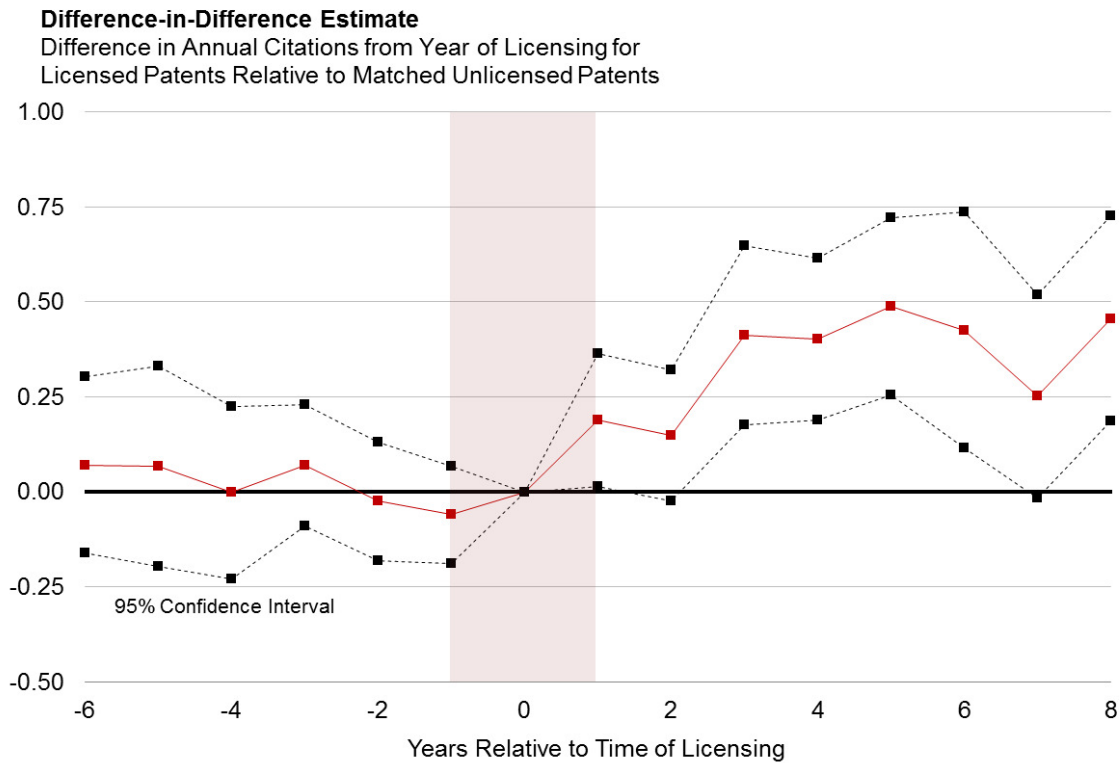
Finally, Model (6) in Table 2.6 extends Model (5) by estimating yearly difference-in-difference effects of licensing. The coefficients presented in this model are each relative to the citation rate in the year the patent (or its matched pair) is licensed. The coefficient on the pre-licensing dummy variables are not distinguishable from zero (except for the year eight years prior to licensing, although this may be an anomaly), indicating that matching on pre-treatment citations worked. In the post-licensing period, the coefficient on the second year after licensing dummy is also not distinguishable from zero. However, beginning in the third year after licensing, the difference-in-difference rate is statistically significant. This effect holds for years three through eight after licensing (although the dummy on seven years after licensing is not statistically significant). During this period, citations are between 0.253 – 0.465 cites/year higher to licensed patents relative to unlicensed patents of comparable age. This represents a 36 – 65% increase in the citation rate for licensed patents due to licensing. Nine to ten years after licensing, the estimated difference-in-difference coefficient drops, but this is likely due to insufficient data on patents that have been licensed for this length of time. The surprisingly negative coefficient on the ten year post-licensing dummy is also the least precisely estimated of the post-period dummies. The results of Model (6) are presented graphically in Figure 2.7.

**Table 2.6: Baseline Regression Estimates**

	(1)	(2)	(3)	(4)	(5)	(6)
	Post-Licensing Cross-Section	Time-Series	Diff-in-Diff	Diff-in-Diff matched pair FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE yearly effects
Dependent Variable Functional Form	All Cites OLS	All Cites OLS	All Cites OLS	All Cites OLS	All Cites OLS	All Cites OLS
Matched Sample	Preferred matching post-license period	Preferred matching licensed patents only	Preferred matching	Preferred matching	Preferred matching	Preferred matching
EVER_LICENSED	0.344 *** (0.100)		0.015 (0.021)	-0.015 (0.015)		
LICENSED		0.520 *** (0.155)	0.328 *** (0.091)	0.335 *** (0.091)	0.223 *** (0.066)	
POST			0.145 (0.110)	-0.046 (0.067)		
PRE_LICENSE(10)						-0.294 (0.249)
PRE_LICENSE(9)						-0.310 (0.269)
PRE_LICENSE(8)						-0.566 * (0.246)
PRE_LICENSE(7)						-0.040 (0.218)
PRE_LICENSE(6)						0.106 (0.126)
PRE_LICENSE(5)						0.049 (0.137)
PRE_LICENSE(4)						0.007 (0.121)
PRE_LICENSE(3)						0.095 (0.084)
PRE_LICENSE(2)						-0.031 (0.080)
PRE_LICENSE(1)						-0.055 (0.066)
POST_LICENSE(1)						0.184 * (0.091)
POST_LICENSE(2)						0.123 (0.090)
POST_LICENSE(3)						0.372 ** (0.123)
POST_LICENSE(4)						0.358 ** (0.114)
POST_LICENSE(5)						0.445 *** (0.126)
POST_LICENSE(6)						0.381 * (0.164)
POST_LICENSE(7)						0.253 (0.140)
POST_LICENSE(8)						0.465 *** (0.140)
POST_LICENSE(9)						0.155 (0.144)
POST_LICENSE(10)						-0.379 * (0.191)
Age Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Matched Pair Fixed Effects?	No	No	No	Yes	No	No
Patent Fixed Effects?	No	No	No	No	Yes	Yes
Observations	6,621	4,712	9,357	9,357	9,357	9,357
Number of Patents	808	404	808	808	808	808
Adj. R-Squared	0.094	0.127	0.096	0.086	0.093	0.102

\* p < 0.05, \*\* p < 0.01, \*\*\*, p < 0.001

Regression equations as described in Section 2.4.2. Models (1) – (6) correspond to Equations (2.1) – (2.6). Standard errors clustered at the patent level in models (1) – (2) and (5) – (6) and clustered at the matched-patent level in models (3) – (4).



Estimated diff-in-diff coefficients associated with individual years before and after licensing. Beginning 2 years after licensing, citations increase by about 0.25 – 0.47 cites/year through 8 years after licensing. Quantitative estimates shown in Model (6) in Table 2.6.

**Figure 2.7:** *Annual Difference-in-Difference Estimates*

I assess sensitivity to functional form in Table 2.7. Models (1) and (2) in Table 2.7 replicate Models (5) and (6) in Table 2.6 for comparison. Models (3) and (4) in Table 2.7 apply a negative binomial functional form to Equations (2.5) and (2.6). Because of restrictions in the negative binomial model, patents with no citations over the panel are dropped, reducing the number of patent-year observations from 9,357 to 7,640. Nevertheless, the estimated coefficients in the negative binomial regressions closely match the OLS results in sign and statistical significance. Models (5) and (6) present results from an augmented log-linear regression where the dependent variable is transformed by the function  $\log(CITES + 1)$  to account for its skewness. Again, results match closely to the OLS results in both sign and statistical significance.

**Table 2.7: Robustness to Functional Form**

	(1)	(2)	(3)	(4)	(5)	(6)
	Diff-in-Diff patent FE	Diff-in-Diff patent FE yearly effects	Diff-in-Diff patent FE	Diff-in-Diff patent FE yearly effects	Diff-in-Diff patent FE	Diff-in-Diff patent FE yearly effects
Dependent Variable	All Cites	All Cites	All Cites	All Cites	ln(All Cites + 1)	ln(All Cites + 1)
Functional Form	OLS	OLS	Negative Binomial	Negative Binomial	OLS Log-Linear	OLS Log-Linear
Matched Sample	Preferred matching	Preferred matching	Preferred matching	Preferred matching	Preferred matching	Preferred matching
LICENSED	0.223 *** (0.066)		0.265 *** (0.068)		0.067 ** (0.022)	
PRE_LICENSE(10)		-0.294 (0.249)		-0.659 (0.455)		-0.180 (0.093)
PRE_LICENSE(9)		-0.310 (0.269)		-0.966 * (0.411)		-0.220 * (0.095)
PRE_LICENSE(8)		-0.566 * (0.246)		-1.018 ** (0.340)		-0.248 ** (0.096)
PRE_LICENSE(7)		-0.040 (0.218)		-0.422 (0.238)		-0.051 (0.079)
PRE_LICENSE(6)		0.106 (0.126)		-0.170 (0.191)		0.012 (0.049)
PRE_LICENSE(5)		0.049 (0.137)		-0.246 (0.174)		-0.026 (0.050)
PRE_LICENSE(4)		0.007 (0.121)		-0.152 (0.155)		-0.034 (0.042)
PRE_LICENSE(3)		0.095 (0.084)		0.036 (0.130)		0.030 (0.035)
PRE_LICENSE(2)		-0.031 (0.080)		-0.076 (0.124)		-0.036 (0.030)
PRE_LICENSE(1)		-0.055 (0.066)		0.001 (0.109)		-0.025 (0.029)
POST_LICENSE(1)		0.184 * (0.091)		0.231 * (0.098)		0.040 (0.029)
POST_LICENSE(2)		0.123 (0.090)		0.176 (0.100)		0.013 (0.029)
POST_LICENSE(3)		0.372 ** (0.123)		0.375 *** (0.099)		0.075 * (0.033)
POST_LICENSE(4)		0.358 ** (0.114)		0.396 *** (0.104)		0.096 ** (0.033)
POST_LICENSE(5)		0.445 *** (0.126)		0.500 *** (0.109)		0.119 *** (0.036)
POST_LICENSE(6)		0.381 * (0.164)		0.486 *** (0.119)		0.104 ** (0.037)
POST_LICENSE(7)		0.253 (0.140)		0.464 *** (0.132)		0.064 (0.044)
POST_LICENSE(8)		0.465 *** (0.140)		0.783 *** (0.134)		0.182 *** (0.047)
POST_LICENSE(9)		0.155 (0.144)		0.605 *** (0.157)		0.097 (0.050)
POST_LICENSE(10)		-0.379 * (0.191)		0.338 (0.195)		-0.043 (0.062)
Age Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Matched Pair Fixed Effects?	No	No	No	No	No	No
Patent Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,357	9,357	7,640	7,640	9,357	9,357
Number of Patents	808	808	615	615	808	808
Log Likelihood			-5,993.1	-5,967.6		
Wald Chi-Squared			839.6	880.4		
Adj. R-Squared	0.093	0.102			0.152	0.160

\* p < 0.05, \*\* p < 0.01, \*\*\*, p < 0.001

Next, I assess sensitivity to the matching algorithm. I rely heavily on matching for causal identification, so assessing sensitivity to different decisions in the matching algorithm is important. Table 2.8 shows results from five matching algorithms applied to estimating Equation (2.5). Figure 2.8 displays these estimates graphically and Figure 2.9 displays the results of the thirteen matching approaches applied to estimating Equation (2.6). Model (2) in the table is the “preferred specification” and replicates Model (5) in Table 2.6. The other models implement matching algorithms as described in Section 2.4.1. Overall, the choice of number of topics does not appear to affect the results as much as choice of what to do with the topics (matching on estimated hazard, Mahalanobis distance, or CEM on topic peaks). Other than the estimate in Model (6) for the 50-topic model with CEM on topic peaks with ties resolved by the estimated hazard, the estimates in models (9) – (13) which rely on exact matching on primary USPTO classification are systematically larger than the other estimates. These estimates also have larger standard errors, most likely due to having dropped a large number of observations to create exact matches on the USPTO classes. These findings are generally consistent with the Rosenbaum bounds approach shown in Table 2.5. Overall, the range of estimates across matching approaches represent a 31 – 48% increase in the citation rate to licensed patents after licensing. The thirteen approaches to matching produce similar estimates, and each approach confirms the overall qualitative thesis of this chapter. In fact, the preferred specification that I focus on is at the lower end of this range of estimates.

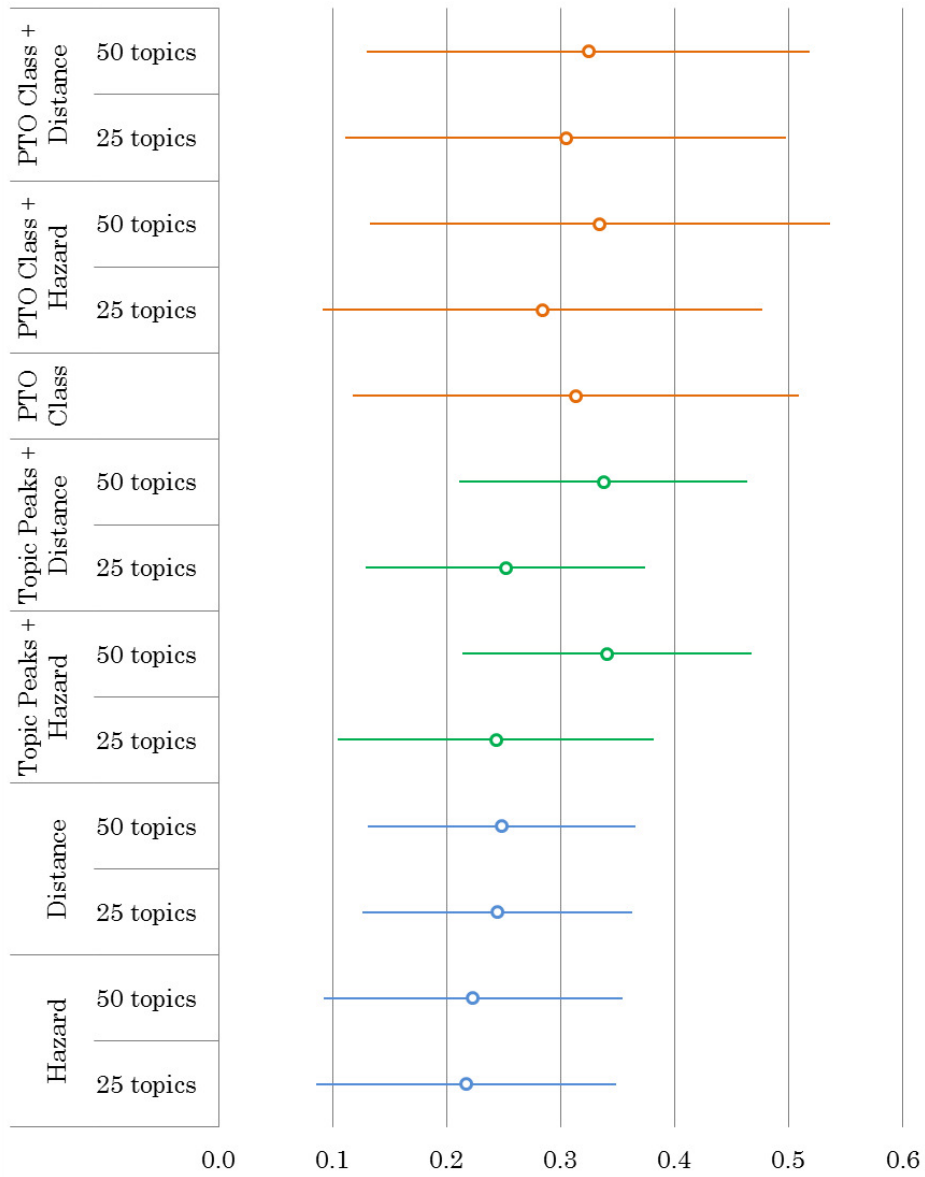
**Table 2.8: Robustness to Matching Design**

	(1)	(2)	(3)	(4)	(5)
	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE
Dependent Variable	All Cites	All Cites	All Cites	All Cites	All Cites
Functional Form	OLS	OLS	OLS	OLS	OLS
Matched Sample	Hazard (25 topics)	Hazard (50 topics)	Distance (25 topics)	Distance (50 topics)	Topic Peaks + Hazard (25 topics)
LICENSED	0.217 *** (0.067)	0.223 *** (0.066)	0.244 *** (0.060)	0.248 *** (0.059)	0.243 *** (0.070)
Age Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Matched Pair Fixed Effects?	No	No	No	No	No
Patent Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Observations	9,377	9,357	9,366	9,441	8,134
Number of Patents	809	808	810	815	709
Adj. R-Squared	0.090	0.093	0.108	0.099	0.090

	(6)	(7)	(8)	(9)	(10)
	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE
Dependent Variable	All Cites	All Cites	All Cites	All Cites	All Cites
Functional Form	OLS	OLS	OLS	OLS	OLS
Matched Sample	Topic Peaks + Hazard (50 topics)	Topic Peaks + Distance (25 topics)	Topic Peaks + Distance (50 topics)	PTO Primary Class	PTO Primary Class + Hazard (25 topics)
LICENSED	0.340 *** (0.064)	0.251 *** (0.062)	0.337 *** (0.064)	0.313 ** (0.099)	0.284 ** (0.097)
Age Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Matched Pair Fixed Effects?	No	No	No	No	No
Patent Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Observations	7,349	8,176	7,431	2,904	2,968
Number of Patents	648	711	655	273	277
Adj. R-Squared	0.099	0.103	0.099	0.097	0.088

	(11)	(12)	(13)
	Diff-in-Diff patent FE	Diff-in-Diff patent FE	Diff-in-Diff patent FE
Dependent Variable	All Cites	All Cites	All Cites
Functional Form	OLS	OLS	OLS
Matched Sample	PTO Primary Class + Hazard (50 topics)	PTO Primary Class + Distance (25 topics)	PTO Primary Class + Distance (50 topics)
LICENSED	0.334 *** (0.102)	0.304 ** (0.098)	0.324 *** (0.098)
Age Fixed Effects?	Yes	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes	Yes
Matched Pair Fixed Effects?	No	No	No
Patent Fixed Effects?	Yes	Yes	Yes
Observations	2,869	2,962	2,979
Number of Patents	268	278	280
Adj. R-Squared	0.087	0.093	0.092

\* p < 0.05, \*\* p < 0.01, \*\*\*, p < 0.001

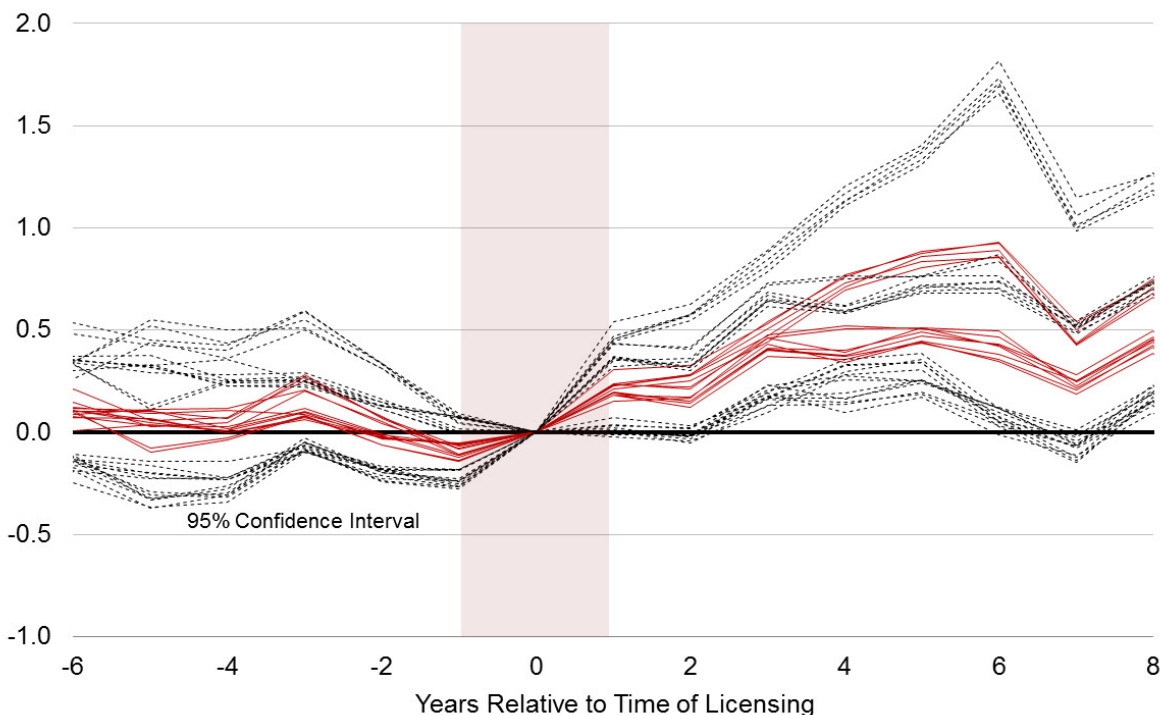


Coefficient Estimate for Diff-in-Diff Annual Citations for Licensed/Unlicensed Patents Before/After Licensing

Estimates and 95% confidence intervals are shown for the matching procedures described in Section 2.4.1 for the regression model described by Equation (2.5).

**Figure 2.8:** *Sensitivity of Coefficient Estimates under Different Matching Procedures*

**Difference-in-Difference Estimates Across Matching Designs**  
 Difference in Annual Citations from Year of Licensing for Licensed  
 Patents Relative to Matched Unlicensed Patents



Estimates (red lines) and 95% confidence intervals (black dashed lines) are shown for the matching procedures described in Section 2.4.1 for the regression model described by Equation (2.6).

**Figure 2.9:** *Sensitivity of Coefficient Estimates under Different Matching Procedures*

## 2.5.2 Exclusive versus Non-Exclusive Licenses

I collected information from the Labs on which patents were exclusively licensed or non-exclusively licensed. In general, Lab technology transfer offices prefer to offer less exclusive licenses so as to increase the number of users who can access their technologies, whereas licensees prefer greater exclusivity so as to protect their right. In my sample, 49% (430 of 877 patents in the full sample) of licensed patents are exclusively licensed. Table 2.9 adds in a separate variable that interacts the LICENSED dummy with a dummy indicating whether the license was on an exclusive basis. The coefficient on the interaction term is positive and suggests a 50% increase in follow-on



innovation relative to non-exclusive licenses. However this difference is not statistically significant, indicating that this difference is not precisely estimated. Because non-exclusive licenses still typically are exclusive within their field of use or within a geographic region, this finding is not completely surprising.

In terms of mechanisms, a license agreement can cause greater follow-on innovation either through a signaling effect or a learning-by-doing effect. A non-exclusive license and exclusive license would seem to suggest an equal signal but provide greater incentives for learning-by-doing, as a firm is likely to have a greater market share under exclusive licensing. This result provides preliminary evidence that the learning-by-doing effect is greater than the signaling effect, as exclusive licensing adds approximately 50% additional citations.

**Table 2.9:** *Exclusive versus Non-Exclusive Licenses*

	(1)	(2)
	Diff-in-Diff	Diff-in-Diff
Dependent Variable	All Cites	All Cites
Functional Form	OLS	OLS
Matched Sample	Preferred matching	Preferred matching
LICENSED	0.223 *** (0.066)	0.179 * (0.080)
EXCLUSIVE_LICENSED		0.091 (0.097)
Age Fixed Effects?	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes
Matched Pair Fixed Effects?	No	No
Patent Fixed Effects?	Yes	Yes
Observations	9,357	9,357
Number of Patents	808	808
Adj. R-Squared	0.093	0.093

\* p < 0.05, \*\* p < 0.01, \*\*\*, p < 0.001

### 2.5.3 Breadth versus Concentration of Spillovers

When follow-on innovation occurs only within a licensing firm, there is no positive externality associated with licensing. Therefore, the broad diffusion of follow-on innovation is important to understand for evaluating the impact of licensing.

Licensing grants a single firm the right to commercialize a patented invention in a field of use. One concern with the empirical findings presented in Section 2.5.1 is that the increased citation rate to licensed patents could be driven by the licensing firm repeatedly developing follow-on inventions from its licensed patent. This would suggest that while licensing leads to follow-on innovation, benefits from induced innovations would be captured by the licensing firm exclusively. To investigate this concern, I create a new dependent variable that counts citations only from patents with assignees who have not already filed a patent citing this same patent. For example, if a patent is cited five times by a single firm, I only count this as one unique citation. Table 2.10 presents results from regressions with this modified dependent variable. Model (1) replicates Model (5) from Table 2.6 for comparison. Model (2) uses the dependent variable of only citations from first-time citers. This model estimates a positive and significant difference-in-difference coefficient, showing that knowledge diffusion from licensing does occur beyond just the licensing firm.

Comparing the magnitudes of the coefficients in Model (1) to Model (2) in Table 2.10 reveals that citations are still concentrated in assignees to a degree. Although I am not able to observe which citing assignees are the licensees, the most conservative interpretation would be to assume that all citations from repeated assignees could be from the licensee. Therefore, comparing these coefficients suggests that at least 76%

(0.169 / 0.223) of the estimated effect in Model (1) is driven by citations from assignees other than the licensing firm.

**Table 2.10: Concentration of Diffusion**

	(1)	(2)
	Diff-in-Diff	Diff-in-Diff
Dependent Variable	All Cites	Cites from Uniquely Assigned Patents
Functional Form	OLS	OLS
Matched Sample	Preferred matching	Preferred matching
LICENSED	0.223 *** (0.066)	0.169 *** (0.042)
Age Fixed Effects?	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes
Matched Pair Fixed Effects?	No	No
Patent Fixed Effects?	Yes	Yes
Observations	9,357	9,357
Number of Patents	808	808
Adj. R-Squared	0.093	0.092

\* p < 0.05, \*\* p < 0.01, \*\*\*, p < 0.001

## 2.5.4 Accounting for Strategic Patenting and Signaling

Sections 2.5.1 – 2.5.3 have established that licensing increases the rate of citation to licensed patents relative to unlicensed patents of comparable age. The increased rate of citations may not necessarily suggest that knowledge is diffusing if citations are accruing to licensed patents for strategic reasons. For example, when a National Lab patent is licensed to a firm, competitor firms may take this as a signal that firms are moving in a certain technological direction. In response these firms could file defensive or “strategic” patents in the area to protect their competitive positions rather than take the information signal as a useful indicator of technological promise, as suggested by Drivas et al. (2014). Strategic patents are for defensive purposes and do not represent

actual knowledge spillovers, yet they still may drive increased citations to the licensed patent – in which case these citations would represent demarcations of prior art rather than inspiration for follow-on innovation. Previous studies have found that strategic patents are cited less often than patents that represent novel inventions (Blind et al., 2009; Harhoff et al., 2003). Therefore, to investigate the mechanisms underlying the results presented previously, I define two new independent variables. First, I create a new variable of citations that only counts citations from patents that themselves have been cited at least once. Second, I create a variable of citations that only counts citations from patents that have received at least the median annual rate of citations. This effectively drops all citation counts in the dependent variable from patents that received fewer than 0.27 citations per year.

Table 2.11 presents the results of regressions applying Equation (2.5) to three dependent variables: all citations, citations only from patents cited at least once, and citations only from patents with at least median citations per year. Model (1) in Table 2.11 again repeats Model (5) in Table 2.6. In my sample, 35% of citing patents are themselves never cited. If never-cited patents were proportionally represented as citers to all patents, then the estimate in Model (2) should be approximately 35% less than the estimate in Model (1). The estimate from Model (2) is 33% less than the estimate from Model (1), suggesting that never-cited patents cite licensed patents approximately just as often as average. A similar comparison for Models (3) and (1) can be made. By definition, 50% of citing patents are cited less than the median of 0.27 cites per year. Therefore, if above-the-median citing patents were evenly distributed, the expected coefficient in Model (3) would be half of the coefficient in Model (1). Table 2.7 shows that the coefficient in Model (6) is indeed 50% smaller than the coefficient in Model (1). Taken

together, these results provide evidence that the additional citations to licensed patents due to having been licensed are from patents that are representative of average citing patents and do not tend to be less often cited themselves. Given that strategic patents are cited less often, this provides evidence against a strong strategic patenting effect.

**Table 2.11: Accounting for Strategic Patenting**

	(1)	(2)	(3)
	Diff-in-Diff	Diff-in-Diff	Diff-in-Diff
Dependent Variable	All Cites	Cites from Non-Zero Cited Patents	Cites from Median Cited Patents
Functional Form	OLS	OLS	OLS
Matched Sample	Preferred matching	Preferred matching	Preferred matching
LICENSED	0.223 *** (0.066)	0.150 *** (0.044)	0.111 ** (0.036)
Age Fixed Effects?	Yes	Yes	Yes
Citing Year Fixed Effects?	Yes	Yes	Yes
Matched Pair Fixed Effects?	No	No	No
Patent Fixed Effects?	Yes	Yes	Yes
Observations	9,357	9,357	9,357
Number of Patents	808	808	808
Adj. R-Squared	0.093	0.101	0.081

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*,  $p < 0.001$

## 2.6 Discussion and Conclusion

Public innovation policy is oriented toward enhancing basic scientific understanding and developing fundamental inventions without immediate commercial application. Yet public R&D has led to new technologies, such as the Internet, GPS, and radar, which have dramatically altered the economy and improved well-being. While initial investment in these inventions relied on public support, development of the majority of these inventions into the products and processes that made these inventions revolutionary required large private investment.

Beginning with the Stevenson-Wydler and Bayh-Dole Acts of 1980, national policy reforms over the past 30 years have greatly increased the intellectual property protection surrounding publicly sponsored inventions, decreasing public access to utilizing these technologies. Yet despite this decreased access, the diffusion of technological knowledge created by public R&D funding has not similarly decreased.

In this chapter, I have shown that in the context of five U.S. National Labs, technology transfer agreements that license patents to private firms have increased the rate of spillovers from publicly sponsored inventions. This empirical finding, corroborated through a variety of statistical methods to account for unobserved heterogeneity and selection bias, provides new evidence for the role of intellectual property protection in increasing the benefits to publicly funded R&D. By making technological knowledge appropriable, patenting allows public institutions to transfer the right to utilize a public invention for the purpose of commercializing the technology in the “market for ideas” (Gans and Stern, 2010). In this chapter, I find evidence that the incentives to commercialize a licensed technology lead to a net positive change in follow-on invention and that this effect is driven not just by the licensing firm, but by invention in firms that did not have access to the licensed technology. Instead these firms may either learn useful information about licensed government patents or may have still gained experience with downstream products and processes resulting from commercialization. This implies that at least for the case of publicly sponsored inventions, the spillover effects of licensing a patent cannot be fully appropriated by the licensing firm.

This research is also provides an important improvement on existing approaches to evaluating technology transfer, which have often relied on short-term and easy-to-

measure metrics. Recent policy evaluation initiatives have called for improved measurement of technology transfer efforts (Stepp et al., 2013; The White House, 2014, 2011; U.S. Government Accountability Office, 2009), and the focus of this chapter on measuring the social impacts of technology transfer through its effect on follow-on innovation would be a useful contribution to this discussion.

Methodologically, the matching algorithms presented in this chapter could be usefully extended to study other problems in the innovation literature. In this work, I have focused on patent abstracts to find comparable patents, but future work could apply text classification to the claims within patents to understand which patents draw on more diverse prior art or establish new technologies that are cross-disciplinary. This would be a particularly useful extension of this mixed-membership model.

Clearly, not all innovations (and not even all important innovations) are captured by patents. While the results presented in this chapter apply to patented inventions, the recurrent problems associated with drawing inferences about innovation more broadly from patent data persist. However, the context for this work is over transactions in the market for innovations, and these markets rely on patents heavily to develop contractible assets. Therefore, the conclusions of this work do have important implications for understanding the general dynamics of innovation and knowledge diffusion.

Licensing a technology to a private firm requires IP protection. In this chapter, I show that licensing an already patented invention increases the rate of knowledge diffusion. The magnitude of the effect I estimate can be usefully compared to other research that has studied the effect of patenting on knowledge diffusion. Such comparisons can provide insight on whether patenting and licensing considered together

positively or negatively effects knowledge diffusion relative to putting a publicly discovered invention into the public domain. Galasso and Schankerman (2013) estimate that removing patent protection on highly valuable patents increases the citation rate to these patents by 50%<sup>25</sup>. The estimates I present in this chapter would suggest that licensing cuts the effect of patenting on knowledge diffusion impediment by more than half. This suggests that for the sole objective of increasing knowledge diffusion, patented inventions should be licensed but unpatented inventions should not be patented because diffusion effects are even higher for inventions in the public domain. However, from the perspective of maximizing the net social benefit of public R&D over time, policy must assess the tradeoff between developing discovered inventions for use in the short-run through greater IP protection with greater knowledge diffusion to create new inventions in the long-run. At the margin, technology transfer is a win-win policy as it both drives private investment into developing technologies in the short-run and also increases the rate of follow-on innovation for the long-run.

## 2.7 Acknowledgments

I would like to thank Laura Diaz-Anadon, Joe Aldy, F.M. Scherer, Bill Clark, Richard Freeman, Venky Narayanamurti, Arthur Spirling, Alberto Abadie, Sam Liss, and Maryanne Fenerjian for their advice at various stages of this project. I also thank seminar participants of the Harvard Environmental Economics Lunch, the Harvard Energy Technology Innovation Policy and Consortium for Energy Policy Seminar, the Harvard Science, Technology and Public Policy Seminar, the Harvard Sustainability

---

<sup>25</sup> Murray and Stern (2007) find that citations to academic papers decline 10-20% when a patent that covers the same invention disclosed in the paper is granted. While citations to papers and patents are very different, the magnitude of the effect is still a useful benchmark.



Science Fellows Seminar, the Global TechMining Conference, the Atlanta Conference on Science and Innovation Policy, and the Technology Transfer Society Conference. I thank Elsie Quaite-Randall, Juliet Hart, Cheryl Fragiadakis, Cheryl Cejka, Matthew Love, Walter Copan, Kimberley Elcess, Jessica Sosenko, Kathryn Klos, Alexandra Andrego, John Lucas, Matthew Ringer, and Sarah Jane Maxted at (or formerly at) the U.S. National Laboratories and U.S. Department of Energy for sharing data and providing helpful comments. All errors are my own.

## Chapter 3

# Financing Wind Energy Deployment in China through the Clean Development Mechanism<sup>26</sup>

In this chapter we study the Kyoto Protocol's Clean Development Mechanism (CDM) using data on all Chinese wind energy projects constructed through the end of 2012. First, we present the first systematic evidence on the pervasiveness of the CDM in the Chinese wind sector. We find that 80% of projects from 2007-2012 received financing through the CDM. As a project-based offsetting mechanism, CDM projects must establish their additionality – that they would not have been constructed without the financial support from the Mechanism – and although additionality claims are verified by a third-party auditing board, calculations depend on several untestable assumptions, such as the expected wind resource quality. In this chapter, we evaluate the claims of

---

<sup>26</sup> Co-authored with Joern Huenteler.

additionality of projects funded by the CDM by calculating the internal rate of return for every project that did and did not receive CDM financing utilizing a state-of-the-art methodology for estimating the distribution of wind speeds at a fine geographic granularity. We find no evidence that CDM projects were any less financially viable than those projects not financed through the CDM, which are non-additional by definition. These results imply that the Chinese wind projects financed through the CDM would have been built in the absence of the CDM, and that the credits generated in this sector, amounting to 3 years of emissions reductions in the European Union Emissions Trading Scheme, did not represent actual emissions reductions.

### **3.1 Introduction**

Anthropogenic climate change is a global commons problem, making international cooperation necessary to significantly mitigate its impacts. Two key features of climate change underpin the necessity of international cooperation: the global mixing of greenhouse gases (GHGs) in the atmosphere and the wide variation in the costs of mitigation opportunities across countries. (IPCC, 2014) These features imply that cost-effective climate change mitigation, which requires equating marginal costs of emissions abatement across all jurisdictions, will mobilize mitigation opportunities wherever they are most affordable. However, the political feasibility of a global policy architecture that directly covers the lowest-cost mitigation opportunities is limited, as these low-cost opportunities are likely to be in the developing countries that bear the least historic responsibility for the accumulated GHGs in the atmosphere. With this motivation, international offsets are seen as a path forward to increase the cost-effectiveness of climate change mitigation while maintaining politically feasible distributional outcomes. International offsets and other flexibility mechanisms can achieve this by geographically

splitting which countries bear the costs of mitigation from which countries actually host mitigation efforts.

The most significant international offset scheme to date has been the Kyoto Protocol's Clean Development Mechanism. Under the CDM, countries with legally binding emission reduction targets under the Kyoto Protocol could meet their targets in part by financing mitigation projects in countries without emission targets. In principle, the CDM enhanced the cost effectiveness and the political feasibility of the Kyoto Protocol by expanding the pool of mitigation opportunities (thus lowering costs), and by creating institutional support for North-South financial transfers. However, the effect of the CDM on the Kyoto Protocol's environmental effectiveness has been called into question due to concerns that credited emission reductions under the CDM were not actually associated with the level of mitigation claimed.

### **3.1.1 Offsets in Theory**

Pollution offsets are credits representing a unit of emissions reduction by an activity outside of a policy's coverage (either outside of the policy's geographic jurisdiction or beyond the policy's covered emission sources). Most prominently, pollution offsets have been implemented as a complimentary policy design feature of quantity based climate policies, such as cap-and-trade schemes.

Offset schemes, such as the CDM, function by creating a pool of alternative credits, referred to as "Certified Emission Reductions" (CERs)<sup>27</sup>, which can substitute for mitigation activities of the policy's covered sources. In a domestic cap and trade scheme, purchased CERs from an approved international or domestic offset scheme can be

---

<sup>27</sup> One CER corresponds to one ton of CO<sub>2</sub>-equivalent in emission reductions.

submitted in lieu of the cap and trade scheme's emission permits. Offset schemes are beneficial for their effect on economic performance and institutional feasibility: they lower aggregate costs of compliance by increasing the supply of mitigation opportunities and they can facilitate the linkage between national climate policies (Ranson and Stavins, 2015). In principle, because of the substitutability enabled by offset schemes for CERs to take the place of domestic pollution reductions, greater utilization of CERs directly implies higher emissions from the sources covered by a domestic policy. With globally mixing pollutants, such as CO<sub>2</sub> and other GHGs, geographic shifting of emission sources and mitigation activities is of little environmental consequence. However, at a fundamental level, whether offset credits represent equivalent emission reductions to the credits that they substitute for cannot be shown incontrovertibly.

Offset credits only represent the real emission reductions they purport to if the credited activity would not have occurred in the absence of the financial transfers associated with the generation of CERs. A version of the world without the transfers of an offset scheme cannot be directly compared to the real world that has an offset scheme, and therefore, whether or not the credited activity was induced by the offset scheme or whether the activity would have happened otherwise, must rely on untestable assumptions about these two worlds. In other words, the environmental integrity of CERs (and therefore an entire offset scheme) depends on whether or not the offsetting project *caused* emissions to decrease. The concept of whether or not an offset project would have occurred only with the offset scheme in place is referred to as "additionality," and lies at the heart of whether or not offset schemes have a negative effect on the environmental effectiveness of the climate policies they complement.

Additionality implies that emission offset projects should be credited for the difference in emissions between the emissions measured in practice and the emissions that would have occurred in the absence of the project, also known as the “baseline.” More generally, the baseline can be thought of as a counterfactual scenario for GHG emissions within a well-defined system’s boundaries. From the well-known Fundamental Problem of Causal Inference (Holland, 1986), counterfactuals are fundamentally unobservable, thereby implying that the calculation at the center of additionality assessment requires untestable assumptions about baselines. Nevertheless, statistical methods to estimate counterfactual baselines are readily available, but require data for comparable “control groups.” In this study, we utilize data from projects constructed during the same time period when the CDM was active to construct a comparable control group.

### **3.1.1.1 Emissions Additionality and Financial Additionality**

Additionality of offsetting activities must satisfy the criterion that the emissions in the world with the offsetting activity are lower than the alternative counterfactual world where the offsetting activity did not take place. For clarity of prose, we refer to additionality in this context as “emissions additionality.” A less stringent form of additionality that is a necessary but not sufficient condition for emissions additionality is what we refer to as “financial additionality.” Under financial additionality, a project is *financed* only in the world with the offsetting scheme in place and not financed in the world without the offsetting scheme, but the effect of the project on emissions in the two worlds is not directly considered. If a project is financially additional, it may or may not

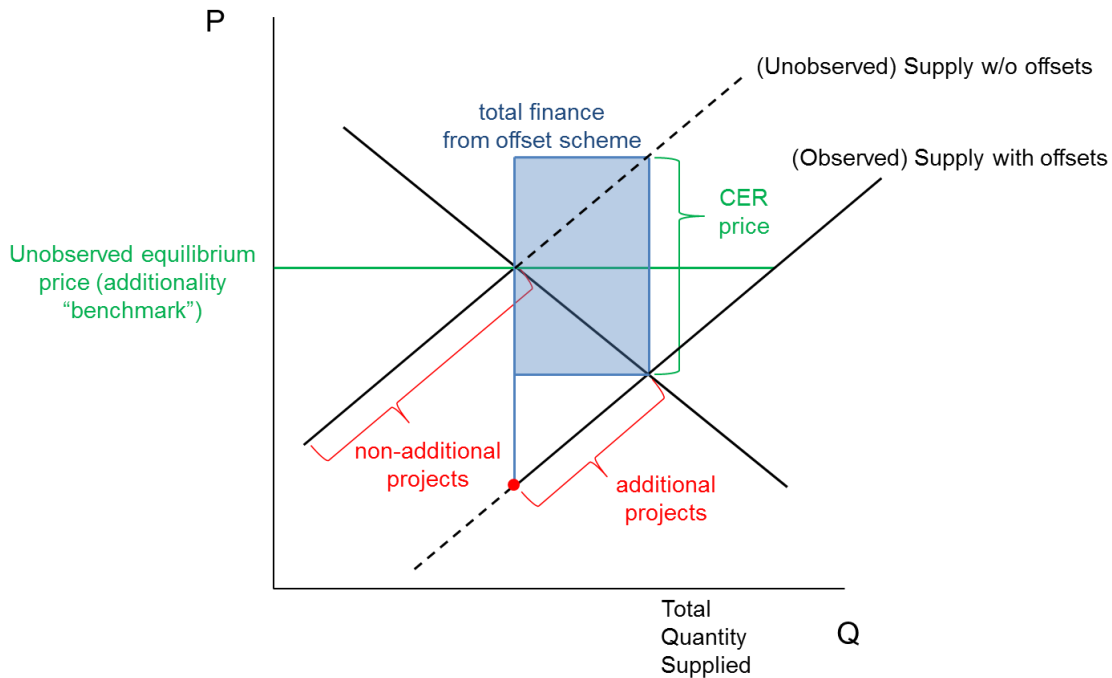
be emissions additional, depending on the project's effect on emissions directly, but also on how the financing of the project affects other activities that produce emissions<sup>28</sup>.

Financial additionality of a project can be assessed with a simple model of an output subsidy. Figure 3.1 displays a schematic diagram of the supply and demand for projects that qualify for offset credits with and without an offset scheme. The presence of the offset scheme acts as a subsidy on the supply of projects and lowers the supply curve. Projects that are financially additional are higher up on the supply curve than projects that are not financially additional. The presence of the offset scheme's effective subsidy on supply increases the quantity of projects by an amount equal to the number of financially additional projects. In practice, assessing financial additionality is challenging because at any point in time, only one supply curve is actually observed and actual marginal costs are typically unobservable. These two challenges manifest themselves clearly in evaluating the financial additionality of projects in practice. Because one supply curve is observed in practice, the determination of financial additionality requires an assumption about the supply curve in the counterfactual world without the offset scheme to determine what the equilibrium point on the supply curve would have been in the absence of the offset scheme. In practice, the equilibrium price in the absence of an offset scheme is referred to as a "benchmark" and is a crucial concept

---

<sup>28</sup> An example of a financially additional but not emissions additional project would be a renewable energy project that received finance through an offset scheme that was not financially viable without this additional finance. However, in response to the construction of this financially additional project, another project developer decided to decommission its (zero-emission) nuclear power project at a one-for-one power production rate. Thus, emissions in the scenario where the offset scheme exists and the renewable energy project is built compared to the scenario in which the offset scheme does not exist and the nuclear power project continues to run would have equal emissions. Similarly, instead of decommissioning the nuclear power plant, complementary infrastructure development to connect the financially additional renewable energy project to the grid may never be made, and thus the project may never generate electricity that can displace other emission-intensive sources, once again making this project financially additional but having no effect on emissions.

in assessing additionality in practice, a point we return to in Section 3.3. This counterfactual equilibrium point can be thought of as determining the financial threshold for assessing whether projects are additional. The second challenge, the opaqueness of marginal costs, is challenging for auditors, particularly when marginal costs rely on highly context-dependent factors (e.g. the wind resource quality for a site where a wind farm is planned).



A single demand curve and a supply curve for a world with an offset scheme and a world without scheme are shown. The offset scheme acts as a subsidy on production and shifts the supply curve downward. Projects that are financially additional have higher marginal costs than non-additional projects, but determining financial additionality in practice is challenging because only one supply curve is observed in practice and marginal costs are typically unobserved.

**Figure 3.1:** *The Supply and Demand for Projects that Qualify for an Offset Scheme*

If a project is financially additional, it may also be emissions additional if the financing of the project also leads to emissions reductions. The assessment of emissions additionality in practice requires assumptions that necessarily vary from project to



project, depending not just on the characteristics of projects themselves but also on the broader context of the projects. Accounting for this broader context is essential in estimating the environmental effectiveness of an offset project because this context informs what other projects might have been constructed in the absence of the project in question. However, accurately modeling a project and its broader context can substantially increase the transactional costs of offset project development.

### **3.1.1.2 Environmental Integrity and Transactions Costs**

Transactions costs associating with verifying offset schemes are increased by challenges of accurately evaluating financial additionality in the absence of reliable cost information and by the observation of supply only in the observed world (and not the counterfactual world without an offset scheme). Together, this sets up a fundamental tradeoff in the design of an offset crediting scheme: offset projects can increase the aggregate cost effectiveness of climate policy by opening the possibility for low-cost mitigation opportunities, but to ensure the environmental integrity of offset projects requires thorough assessment of financial and emissions additionality, increasing transactions costs and lowering the cost effectiveness benefits offsetting schemes might otherwise confer.

This tradeoff is illustrated well by the comparison of project-based offsetting schemes, such as the CDM, and sectoral offsetting schemes, such as those proposed by the European Union for the post-2020 period (UNFCCC, 2012a). Project-based offsetting schemes require every offsetting investment to establish its own financial and emissions additionality, raising transactions costs. In the CDM, while methodologies to facilitate the demonstration of additionality have proliferated, transactions costs remain high (Michaelowa, 2012; Michaelowa et al., 2003). In many cases, financial additionality is

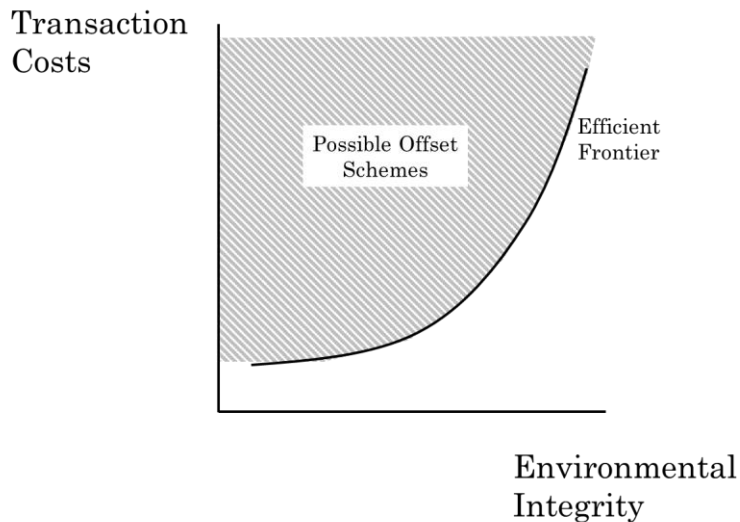
deemed as a sufficient condition for emissions additionality. This is the case for the sector we study, renewable energy. Large-scale projects, including solar photovoltaic technologies, solar thermal electricity generation, off-shore wind technologies, marine wave technologies and marine tidal technologies can follow a “simplified procedure to demonstrate additionality” (UNFCCC, 2014, 2012b). According to the simplified procedure, projects have to demonstrate that (1) the proposed CDM project activity is unlikely to be “financially attractive” (defined by a benchmark) and that (2) there are essential distinctions between the proposed CDM project activity and similar activities that can “reasonably be explained.”

In contrast, sectoral offsetting schemes provide a uniform set of criteria for an entire national sector and offer discounted credits for each project with the assumption that a fraction of projects are not individually additional. The uniform criterion for the sector lowers transactions costs while crediting some projects that are not additional.

The large majority of offset schemes implemented in practice, including the CDM, have been project-based. This choice has been made under the implicit assumption that project-based design of offset schemes is necessary to achieve an acceptable level of additionality, and that the associated higher transactions costs of project-based offset certification and verification are a necessary cost. In other words, high transactions costs are a concession necessary to achieve additionality and therefore the environmental integrity of the original policy.

The conceptual tradeoff between greater environmental integrity of offsets and increasing transactions costs is displayed graphically in Figure 3.2. This tradeoff is essential in prospectively understanding alternative design choices of offset schemes. However, it is also important to understand how existing offset programs have

performed on these dimensions ex-post. In Section 3.3, we present an analysis that accomplishes exactly this for the environmental effectiveness dimension of one of the largest national sectors of the CDM, the Chinese wind sector. In Section 3.3.5, we will revisit how our assessment of the CDM in the Chinese wind sector can inform future policy design of offsetting schemes in light of the tradeoff between environmental effectiveness (a function of additionality) and cost effectiveness (a function of transactions costs, given the available mitigation opportunities).



The conceptual tradeoff between lower transactions costs and higher environmental integrity is primarily due to the analytical complexity of accurately estimating baseline emission scenarios that account for the context that a project is embedded in. The figure makes the implicit assumption that there are decreasing marginal returns in environmental effectiveness from greater transactions costs (implying that further additionality oversight is less productive as a given level of environmental integrity is already ensured). Optimal offset schemes exist on the frontier depicted in the figure, but in practice may lie to the left of the curve, for example because of institutional constraints.

**Figure 3.2:** *Transaction Costs and Environmental Integrity in Offset Schemes*

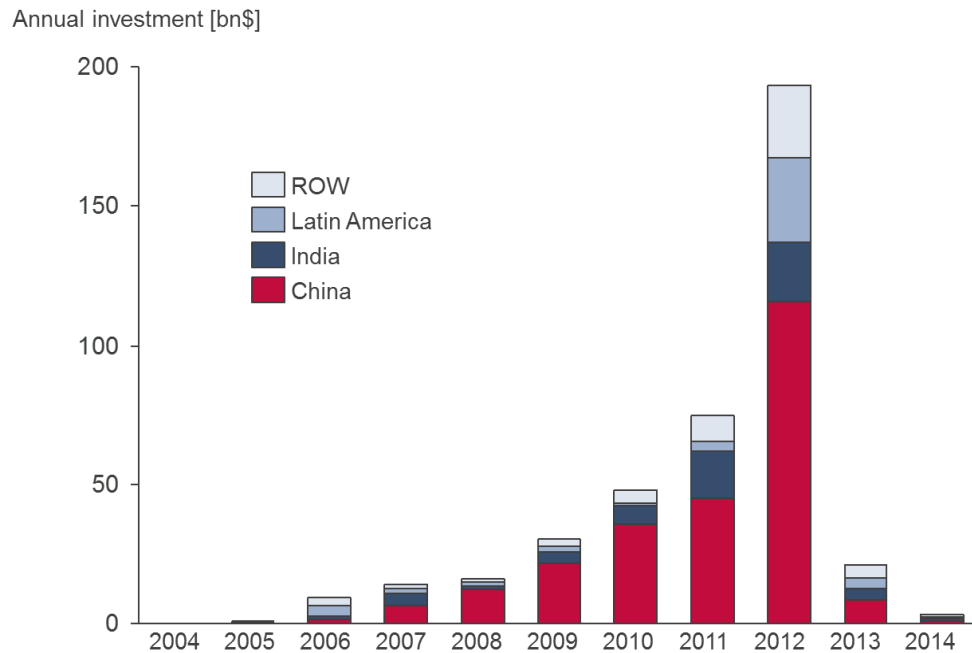
### 3.1.2 Offsets in Practice: The Clean Development Mechanism

International offsets were a key design feature of the 1997 Kyoto Protocol, a legally binding treaty under the United Nations Framework Convention on Climate Change

(UNFCCC). The Kyoto Protocol implemented legally binding pollution targets for largely rich, industrialized countries (Annex B parties) and did not implement binding targets for all other countries. Under the Kyoto Protocol, flexible mechanisms complemented the legally binding emission targets, allowing countries with targets to meet their targets by trading emission permits among themselves, financing additional mitigation projects within countries with obligations (Joint Implementation), or by financing additional mitigation projects in non-Annex B countries (CDM).

As of October 2013, 1.4 billion tons of CO<sub>2</sub>-equivalent in emission permits were created through the CDM. This is equivalent to over two years of global growth in energy-related CO<sub>2</sub> emissions (using the average rate of change for 2000 – 2012) (IEA, 2014). While CDM projects are eligible in virtually all Annex B countries, in practice, CDM projects have been heavily geographically concentrated, as shown in Figure 3.3.

Within China, CDM projects have been concentrated in a small number of sectors. Projects to destroy hydrofluorocarbon gases (HFCs) and other potent industrial gases, are projected to receive 15% of total Chinese CERs and have attracted significant scrutiny for their alleged manipulation of baselines (The Economist, 2010). Renewable energy in the form of hydroelectric and wind energy is expected to receive 55% of Chinese CERs, roughly evenly split between hydroelectric projects (30% of total) and wind energy projects (25% of total). In this chapter, we focus on the 25% of Chinese CERs to be credited to wind energy projects. The distribution of Chinese CDM projects weighted by expected CERs is shown in Figure 3.4.



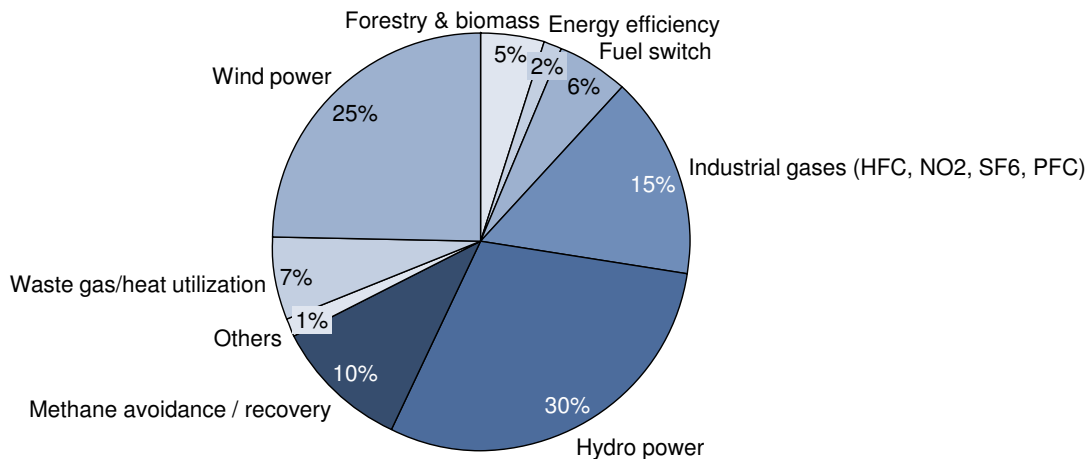
China has hosted 59.02% of all CDM projects, weighted by number of carbon credits issued, more than all other countries combined. Outside of China, India, and Latin America, the rest of the world (ROW) has received 15.39% of the CDM's total carbon credits. As there is a global market for carbon credits, the geographic distribution of issued carbon credits stratified by host country closely mirrors the below distribution of CDM investment.

**Figure 3.3: Geographic Distribution of Investment in CDM Projects**

China's wind sector has benefited significantly from international climate finance through the CDM. And as a result, the case of China's wind sector is often used as a reference to inform the design of offsetting mechanisms and international climate policy more broadly (He and Morse, 2010; Lema and Lema, 2013; Lewis, 2010a; Tang and Popp, 2014).

Beyond the Chinese wind energy sector's importance in its own right as the largest national wind sector in the world, this sector also makes for an interesting context for assessment of additionality because while CDM activity occurred, wind energy was developed simultaneously without support of the CDM, creating a relevant control group of projects. Further, unlike HFC and hydroelectric CDM projects in which CERs are

credited to a very small number of large projects, there are nearly 1,500 approved Chinese wind projects financed by the CDM. The role of the CDM in the context of the Chinese wind sector is explored in Section 3.2 and additionality of these projects is assessed in Section 3.3.



The distribution of Chinese CDM projects by project class weighted by cumulative CERs projected to accrue by 2030. Destruction of hydrofluorocarbon gases (HFCs) is the largest class of Chinese CDM projects, followed by hydroelectric power, wind power, nitrous oxide (N<sub>2</sub>O) emission reduction, energy efficiency, switching away from fossil fuels to other fuels, and capture of methane released from coal beds. (IGES, 2015a)

**Figure 3.4:** *Distribution of Chinese CDM Projects by Project Class*

### 3.1.3 The Chinese Wind Sector

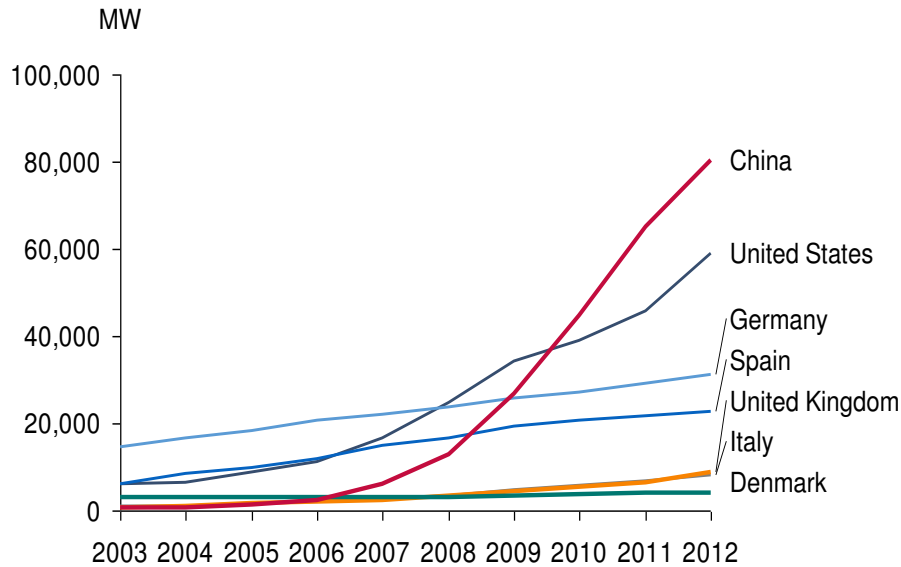
Climate policy in China and the Chinese wind sector have several distinct characteristics that make its study important in and of itself. First, China is currently the world's largest national emitter of GHGs, making efforts to reduce Chinese GHG emissions of global policy relevance. With overtures and examples of several national and sub-national climate policies, China seems poised to expand its GHG mitigation efforts, and renewable energy and offset schemes are likely to play an expanded role moving forward. Second, the Chinese wind sector has experienced exponential growth,

and is now the world's largest national wind sector by installed capacity, with 115 GW of installed capacity at the end of 2014 (NEA, 2015), equivalent to approximately 20% of the world's total installed wind capacity, as shown in Figure 3.5. While becoming the world's largest wind sector, China has also developed a viable domestic industry in turbine manufacturing and wind power services and is now beginning to export turbines to other developing countries (Lewis, 2013).

This rapid growth in the Chinese wind sector was facilitated by several targeted policies, beginning with the Sixth Five-Year Plan in 1981-1985 (Shi, 2003). Through the mid-1990s, Chinese policies were targeted towards small domestically produced turbines and a small number of imported turbine models from Denmark. By 1994, these efforts amounted to approximately 10 megawatts (MW) of installed capacity, equivalent to the capacity of four modern turbines, which can routinely surpass 2.5 MW in capacity. A landmark regulation was adopted in 1994 by the Chinese Ministry of Power, which required electric grid operators to compensate wind power producers with a price of electricity that allowed them to generate a "reasonable profit." Further policies and measures were adopted the late-1990s and early-2000's, including preferential loan schemes (in 1995, 1999, and 2000), national installation targets (in 1996, 1997, and 2001), and support schemes for wind turbine science, technology, and manufacturing (in 1995, 1996, 1997, and 2001). A full list of Chinese policies to facilitate the development of the domestic wind energy sector are described in Appendix C.2 (see also Lewis (2013) and Gosens and Lu (2013)).

Several additional policies from 2003 onward led to even more rapid wind energy development. A new electricity pricing policy in 2003 changed the tendering scheme for large wind parks, tariffs were reformed for all wind projects in 2006, 2008, and 2009, the

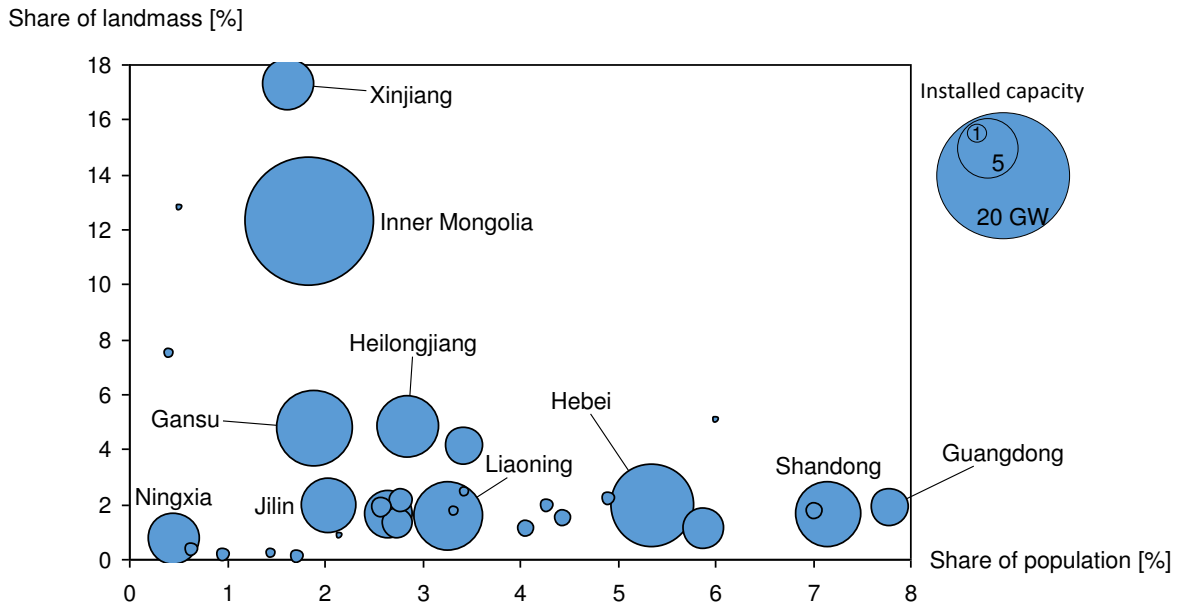
national Renewable Energy Law was adopted in 2005, Chinese wind projects became eligible for CDM project finance following the Kyoto Protocol’s ratification in 2005, and national installation targets were revised upward again in 2006, 2008, and 2009. From 2004 – 2013, a span of nine years, the domestic Chinese wind market grew by 12,800%, equivalent to doubling seven times. Figure 3.5 shows the growth of the Chinese wind market relative to the growth of domestic industries in other countries. Figure 3.6 displays how this growth was distributed across the provinces of China. Importantly for our analysis in Section 3.3, there was no systematic difference in the application of these policies and programs for projects that participated in the CDM and those that did not.



Time series of installed wind capacity for 2003 – 2012. Since 2010, China has had the largest installed capacity of wind power, although the United States currently generates more electricity from wind than China.

**Figure 3.5:** *Cumulative Installed Capacity of Wind Power by Country*





Spatial distribution of Chinese wind farms by province shown by provincial land area and share of population. The area of the bubbles represents the province's total installed capacity as of 2012.

**Figure 3.6:** *Spatial Distribution of Chinese Wind Farms by Province*

The rest of this chapter proceeds as follows. In Section 3.2, we investigate the role of the CDM in facilitating the growth of the Chinese wind sector. In this section, we create the first project-level distinction of CDM projects and non-CDM projects. Prior literature has only examined these two populations of projects in aggregate or individually and has not taken the additional step of matching databases on the universe of Chinese wind projects and databases on the universe of CDM projects. Our distinction allows us to directly investigate the role of the CDM on the projects that received the CDM relative to those projects that did not participate in the CDM. In Section 3.3, we use this project-level matching to assess the claims of additionality of the CDM projects. By definition, projects not supported by the CDM are non-additional, so we can construct careful comparisons of CDM and non-CDM projects to empirically assess the similarity of these

two classes of projects to shed light on counterfactual development in the absence of the CDM.

## **3.2 The Role of the CDM in Financing Wind Energy in China**

### **3.2.1 Introduction**

China's wind power sector is one of the often-cited success stories of low-carbon energy development in emerging economies (Dai and Xue, 2015; Lewis, 2013). However, there remains considerable uncertainty about the exact role played by the CDM and the large amount of international climate finance that flowed to the sector during the period of rapid expansion.

First and foremost, it is unclear what share of China's wind investments received financial support through the CDM. Estimates in the literature of this share range from 32% (He and Morse, 2013) to almost 88% (Stua, 2013). With such a wide range of estimates, it is difficult to assess how fast China's wind sector would have developed without support from the CDM. Second, it remains unclear how important the prospective revenue from offset credits was for getting the projects financed. And third, we know only little about the impact of the CDM on technology choices by investors and technology development more broadly in China's wind turbine industry (Lema and Lema, 2013). Therefore, it is difficult to assess the importance of the CDM in comparison to domestic policy support implemented by the Chinese government. This uncertainty is unfortunate, because a better understanding of the drivers of investment in low-carbon energy in emerging economies is highly relevant for the ongoing debate on the design of the future international climate policy regime, which is likely to have a major technology component, leading up to the envisioned agreement in Paris in December 2015.

This section aims to address this gap by identifying the role of the CDM in attracting investment in China's wind sector. How much of China's wind capacity has been registered to receive offset credits under the CDM? What was the total volume of attracted investment? What was the projected inflow of carbon revenues to the projects that participated in the CDM, and how much revenue has actually been collected so far? And do the projects that received CDM funding use more advanced technology than those that did not?

Our analysis contributes to the political and academic debate on the CDM in two ways. First, our analysis allows us to evaluate the relative importance of domestic and international climate finance in the development of the Chinese wind sector. This helps put in context the lessons drawn for domestic and international climate policy from previous studies of China's wind sector, and, more broadly, has implications for the debate on the future role of offsetting in addressing climate change. Second, our analysis allows us to evaluate the impact of the CDM – in particular its impact on the regional distribution of wind power investments and its impact on technology choice. This contributes to the literature on technology spillovers and international technology transfer. Finally, the analysis in this section allows us to evaluate the environmental effectiveness of the Kyoto Protocol, an analytic effort we pursue in Section 3.3.

The remainder of the section is organized as follows. We first introduce the empirical strategy, data, and methodology in the context of the prior literature in Section 3.2.2. Section 3.2.3 presents our results. Section 3.2.4 discusses the findings, before Section 3.2.5 synthesizes our contributions and the limitations of our analysis.

### 3.2.2 Estimating the Share of CDM Projects in China's Wind Sector

At first glance, the disagreement in the literature about the role of the CDM in China's wind sector might seem surprising. Statistics on total wind installations in China are available from a number of government agencies and industry associations (CREIA, 2014; CREIA et al., 2014; NBS, 2013). And data on each Chinese wind project that receives financial support through the CDM, including size, location, technology, capital investment and investors, is collected by the Secretariat of the United Nations Framework Convention on Climate Change (UNFCCC) and compiled by several institutions (IGES, 2015a, 2015b; UNEP Risoe Center, 2013).

However, integrating and comparing the two data sources is difficult. Chinese official statistics are only consistent for actual installations, not planned projects, while the UNFCCC statistics are consistent only for planned projects, not actual installations. In particular, while the UNFCCC data contains a clear timeline for each project from the start of the planning phase to the end of the project lifetime, it does not keep track of when projects are actually installed. It only indicates when projects *plan* to start generating credits, not when they actual do.<sup>29</sup>

Since some projects may be developed faster than scheduled and others face delays in permitting or construction, using the UNFCCC database to estimate wind power installations under the CDM – as done by several studies (Lewis, 2010a; Stua, 2013; Tang and Popp, 2014) – yields only rough estimates of how many wind projects received CDM support and how many were built outside the CDM. He and Morse (2013) estimate

---

<sup>29</sup> *Actual* operation is only registered when the first monitoring report is submitted, which may be several years after installation. The monitoring reports contain externally verified information on power generation of CDM wind projects and provide the basis for CER issuance.

that about 32% of Chinese wind investments in 2003-2009 received CDM support. This implies that about 17.3 GW of projects were built outside of the CDM in that period. Stua (2013), on the other hand, concludes that about 88% of the 75.56 GW installed by the end of 2012 received credits under the CDM, which would mean that only about 8.6 GW were realized outside the CDM in 2003-2012. Other studies, including by Tang and Popp (2014), who estimate the CDM share at 75% in 2003-2009, Lewis (2010a) (41% in 2003-2008) and Zhang and Wang (2011) (45% in 2005-2008) lie somewhere in between these two extremes. The wide range of estimates, summarized in Table 3.1, poses a significant challenge to research aiming to assess the role of the CDM in the development of the Chinese wind sector.

**Table 3.1:** *Estimates of the Share of Chinese Wind Projects that Received CDM<sup>30</sup>*

Study	Time period	Total capacity installations in period*	CDM Share	Capacity in CDM	Capacity outside CDM
He and Morse (2013)	2003-2009	25.25 GW	32%	8.11 GW	17.23 GW
Lewis (2010)	2003-2008	11.47 GW	41%	6.86 GW	4.69 GW
Zhang and Wang (2011)	2005-2008	10.75 GW	45%	7.18 GW	4.84 GW
Tang and Popp (2013)	2003-2009	25.25 GW	75%	18.94 GW	6.31 GW
Stua (2013)	Until 2012	75.33 GW	88%	66.29 GW	9.04 GW

Our study resolves this uncertainty by individually matching data on CDM and non-CDM projects. We do so by matching two separate, project-level databases, one on

---

<sup>30</sup> The absolute capacity data listed in the table were derived by applying percentage estimates to capacity data provided by Chinese Renewable Energy Industry Association (CREIA et al., 2014) for reasons of consistency and may therefore differ slightly from absolute numbers cited in individual studies.

Chinese CDM wind projects and one on all Chinese wind projects. The comparison of the two matched databases allows us to identify which projects received CDM funding and which did not, and to compare the characteristics of the two populations of projects. In contrast to previous, sector-level studies, the richness of project-level data allows to draw novel and much more rigorous conclusions on the role of the CDM in the development of China's wind sector.

### 3.2.2.1 Data

Our analysis is based on three main databases. *First*, we obtained, cleaned, and augmented a proprietary project-level database on Chinese wind installations (Huaxia, 2013) (from here on referred to as “full-sector database”). The database contains information on each batch of turbines (identified by a unique turbine model; e.g., “33 1.5MW Goldwind turbines”) installed in China at the end of 2012, including the name of the wind farm (in Chinese), the turbine owner, the turbine model and the installation date<sup>31</sup>. In total, the database contains 2,246 batches of turbines with 80.474 GW of capacity<sup>32</sup>. We cleaned the names of turbine models and owners (e.g., by combining subsidiaries and parent companies) and added approximate project coordinates based on the project name (which typically contains the township or village name).

*Second*, we compiled a database of all wind projects in China that were registered with the CDM as of January 2014 (henceforth referred to as the “CDM database”). The

---

<sup>31</sup> The database does not contain information on decommissioned turbines, but a comparison with project-level data from 2002 (Shi, 2003) indicates that the number of decommissioned turbines is very low (99.3% of the turbines installed in 2002 were still online at the end of 2012).

<sup>32</sup> The installed capacity according to Huaxia's methodology is 6.8% higher than data by the Chinese Renewable Energy Industry Association (CREIA, 2014) because it applies a more “generous” definition to the installation date: i.e., turbines are included earlier than in the CREIA data. The installed capacity at the end of 2011 is the same in the two databases (62.4 GW), which indicates that the commissioning dates are adjusted ex-post.

publicly available CDM databases (IGES, 2015a, 2015b; UNEP Risoe Center, 2013) contain information on project timeline, projected CER and power generation, as well as financial information (e.g., investment cost, power tariffs, CER-price assumptions, etc.), which we cleaned and consolidated. To this, we added data on installed turbine models, project owner information, and project coordinate data, which we extracted from the individual Project Design Documents (PDDs)<sup>33</sup>. The final CDM database contains 2,051 batches of turbines (1,494 CDM projects with up to 7 batches each and a total of 82.798 GW).

*Third*, we consolidated and augmented the information on wind turbine models in the two database discussed above using a database on wind turbine models employed in China (Huaxia, 2011). The resulting turbine database contains information on turbine size, the source of the technology (e.g., import, license, local development) and the type of company (e.g., foreign manufacturer, local subsidiary of foreign manufacturer, joint venture, domestic manufacturer, etc.) for 424 unique turbine models (identified by turbine size in kW and manufacturer), allowing us to assess the influence of the CDM on technology choice and technology sourcing. The three categories of technology sourcing used in this chapter, *hardware import*, *locally owned IP*, and *locally developed IP* are defined in Table 3.2.

---

<sup>33</sup> The PDDs are available on the UNFCCC website: <https://cdm.unfccc.int/Projects/projsearch.html>.

**Table 3.2: Technology Transfer Classifications**

Mode	Definition
Hardware import	We define “hardware import” as cases where wind turbines were directly imported from abroad. In case a multinational company established a manufacturing subsidiary in China, only turbines supplied before the establishment of its local subsidiary are defined as imported hardware.
Locally owned IP	A turbine manufactured by a locally owned company that has acquired IP (e.g., by ordering a turbine design from an engineering design company) or has developed its own IP (see below). Turbine licenses were excluded from this definition for the purpose of this analysis. Also, for practical purposes, joint ventures with foreign companies are excluded from this definition independent of majority ownership in this chapter.
Locally developed IP	A turbine manufactured by a locally owned company with its own innovation and proprietary technology design. Imported components may be used in the manufacturing process. But it is considered a local technology as long as the technology design originates in a local company, with or without local technology transfer.

### 3.2.2.2 Database Matching

To make the two datasets comparable, we manually matched the projects in the CDM database with the projects in the full-sector database. Our matching procedure followed a four-step process. *First*, we generated a set of candidate matches that (i) utilized the same turbine model and (ii) were installed in the same province. We then identified matches using additional information on the size, name, project owner, and exact location (township or village) of the wind farm<sup>34</sup>. This step was repeated several times after eliminating identified matches from the pool of candidates and rerunning the matching algorithm. *Second*, we tested for false positives by automatically identifying projects that have (a) a distance of more than 50 km or (b) an installation date in the full-sector database that does not fall in between the “starting date of the CDM project activity” and the “starting date of the crediting period” in the CDM database. We manually checked all UNFCCC project documents of the remaining unmatched 354

---

<sup>34</sup> In some cases, a project in either of the two databases had to be split into multiple batches when all other indicators suggested a match but the total size of the wind farm did not match (e.g., 33 1.5 MW turbines in database A, 30 1.5 MW turbines in database B).



CDM projects for additional information on project location, timeline, and project ownership, and compared this information to all candidates in the same province in the full-sector database. In most cases, we were able to resolve matching conflicts through this additional research, e.g., because information in the publicly available CDM databases did not correspond to the project documents, the turbine model had been changed in later stages of the project cycle, or because coordinate data obtained from the PDDs was incorrect. *Third*, we tested for false positives by checking which of the CDM projects for which we had not found a match in the full-sector database had submitted monitoring reports to the UNFCCC by November 2014 (indicating that they had been built). Here, too, we were able to resolve most conflicts.

The remaining false-negatives error can be estimated at 4%, based on the ratio of the number of CERs issued to all Chinese wind projects (122.6 million CERs) to the number of CERs issued to the matched projects (117.7 million) according to UNFCCC data (IGES, 2015a).

Overall, we matched 78% of the CDM projects in capacity terms and 80% of the full-sector projects. The final, integrated database contains 1,586 matched batches. 565 CDM batches remain unmatched. 256 have a starting date before the end of 2012, which suggests that their implementation has been delayed until after 2012 (the other 409 have a starting date after 2012, the end of the full-sector database). 660 full-sector batches do not have a matching CDM project (562 of which were installed since 2003).

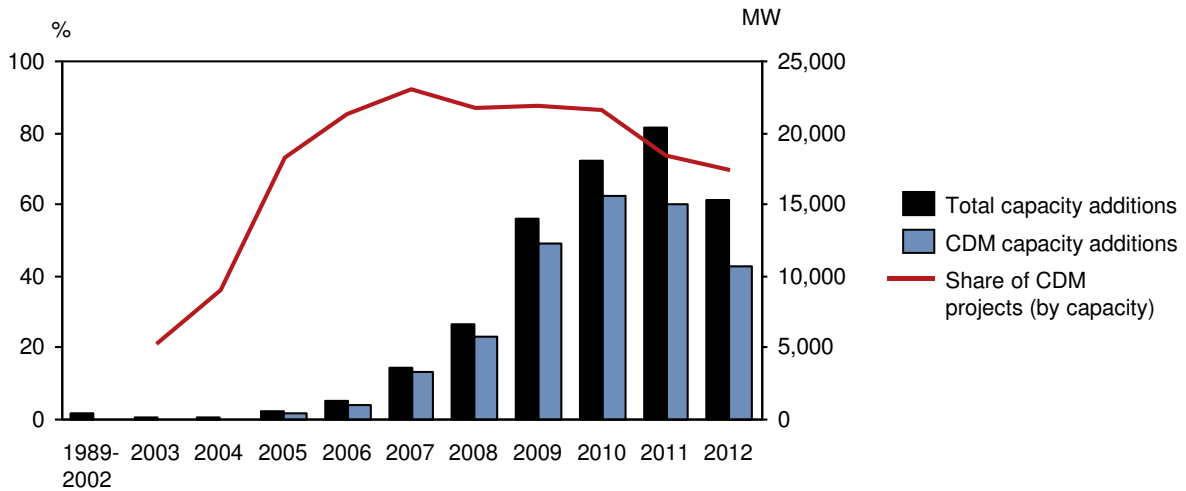
### **3.2.3 Results**

We present four sets of results. First, we show aggregate statistics for the two databases, including the share of Chinese wind projects that received CDM funding. Second, we make use of the financial data in the CDM project documentation to estimate

the financial impact of the CDM on the Chinese wind sector, including the attracted capital investment and the revenues from CER sales. Third, we show a comparison of turbine model choices in CDM and non-CDM projects to illustrate the CDM's impact on technology choice and technology sourcing. And fourth, we show data for the geographical proximity of investments to previous projects to illustrate whether or not the CDM incentivized investments in new and potentially more risky areas.

### **3.2.3.1 Capital Investment under the CDM**

Our results indicate that the CDM provided financial support for a large majority of wind projects installed in China. Figure 3.7 shows capacity additions inside and outside the CDM in 2003, the year of the first CDM installation in China, through 2012. Over this time period, a total of 64.01 GW was installed under the CDM, corresponding to 79.96% of total installations in 2003-2012 and 79.50 % of total cumulative installations since 1989. These numbers are close to the highest estimate in the literature (Stua, 2013). The annual share of CDM projects was lowest in 2003 (21%), when projects could apply for the CDM but the CDM had not gone into full effect, and highest in 2007 (92%). Our result for 2006-2010 are higher than almost all other annual estimates provided in the literature (Lewis, 2010a; Tang and Popp, 2014).



Installed wind capacity in China (total and under the CDM) and share of projects that have received CDM funding over time.

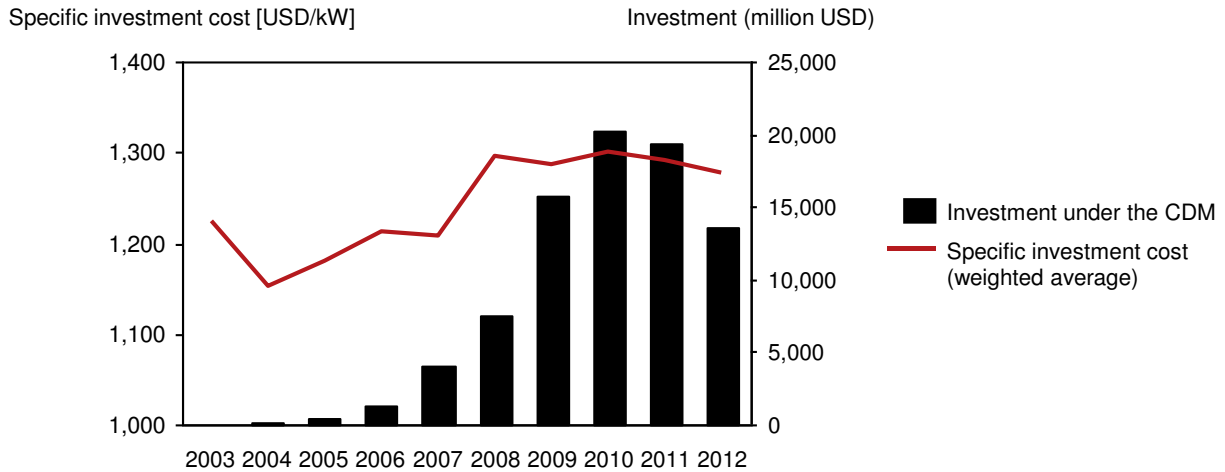
**Figure 3.7:** *Installed Wind Capacity in China under the CDM*

The capital investment attracted for Chinese wind CDM projects in the decade 2003-2012 was substantial. Figure 3.8 shows annual capital investment flows as documented in the project design and validation documents (IGES, 2015b)<sup>35</sup>. Total investment was \$82.21 billion, with annual flows exceeding \$5 billion every year since 2008. Average investment per kW of rated turbine capacity increased over time from below \$1,152 per kW in 2004 to \$1,300 per kW in 2010<sup>36</sup>.

<sup>35</sup> In 116 cases no investment cost data was available in the IGES (IGES, 2015b) database. In these cases we interpolated capital investment using year, province, turbine size, and turbine manufacturer as determinants.

<sup>36</sup> This wind-farm cost increase is similar to trends observed in the United States, where increases have been linked, in part, to increasing turbine complexity and commodity prices (Bolinger and Wiser, 2012). In the Chinese case, the price increase may also be related to the exploitation of increasingly remote wind resources in the Northwestern and Northeastern provinces.

However, our results also show that projects totaling 16.503 GW, corresponding to about \$21.17 billion of investment, were realized in China without the prospect of revenues from the CDM<sup>37</sup>.



Capital investment and revenues are shown for 2003 – 2013 in nominal terms.

**Figure 3.8:** *Capital Investment under the CDM in the Chinese Wind Sector*

### 3.2.3.2 Revenues from the CDM

The size of capital investment under the CDM displayed in Figure 3.8 is in stark contrast to the relatively small flow of revenues from offset credit sales to Chinese wind projects since 2003. On the one hand, the projected revenues from offset credit sales are large in absolute terms but small compared to revenues from electricity sales by the projects. On the other hand, actual CER generation has so far fallen far short of projected CER generation.

Subpanel (a) of Figure 3.9 shows projected emission reductions in tons of CO<sub>2</sub>-equivalent – equal to the projected generation of CDM offset credits, or CERs – by the

---

<sup>37</sup> The investment cost per project were approximated from the CDM investment data based on the following characteristics: province, year, turbine model, and turbine manufacturer.

matched CDM batches over their project lifetime (CDM projects can generate CERs for a maximum of 3 crediting periods of 7 years each, i.e., for up to 21 years). At its peak between 2014 and 2021, the 1,586 projects were expected to generate 133 million CERs per year, with a cumulative total of 2.76 billion tons of avoided emissions over the period 2003-2035. However, actual revenues have fall far short of this projection since 2010: CER issuance reached a peak of 36 million for the year 2011 but fell to 33 million and 5 million in 2012 and 2013, respectively, reaching a cumulative total of 117.7 million in March 2015. Some delay is inherent in the process<sup>38</sup>, but new issuance has almost come to a halt, with a mere 8.46 million in new CERs issued for all 1,586 matched CDM batches between January 31, 2014 and March 30, 2015 (IGES, 2015a), compared to 55 million CERs issued in the 16 month prior to that.

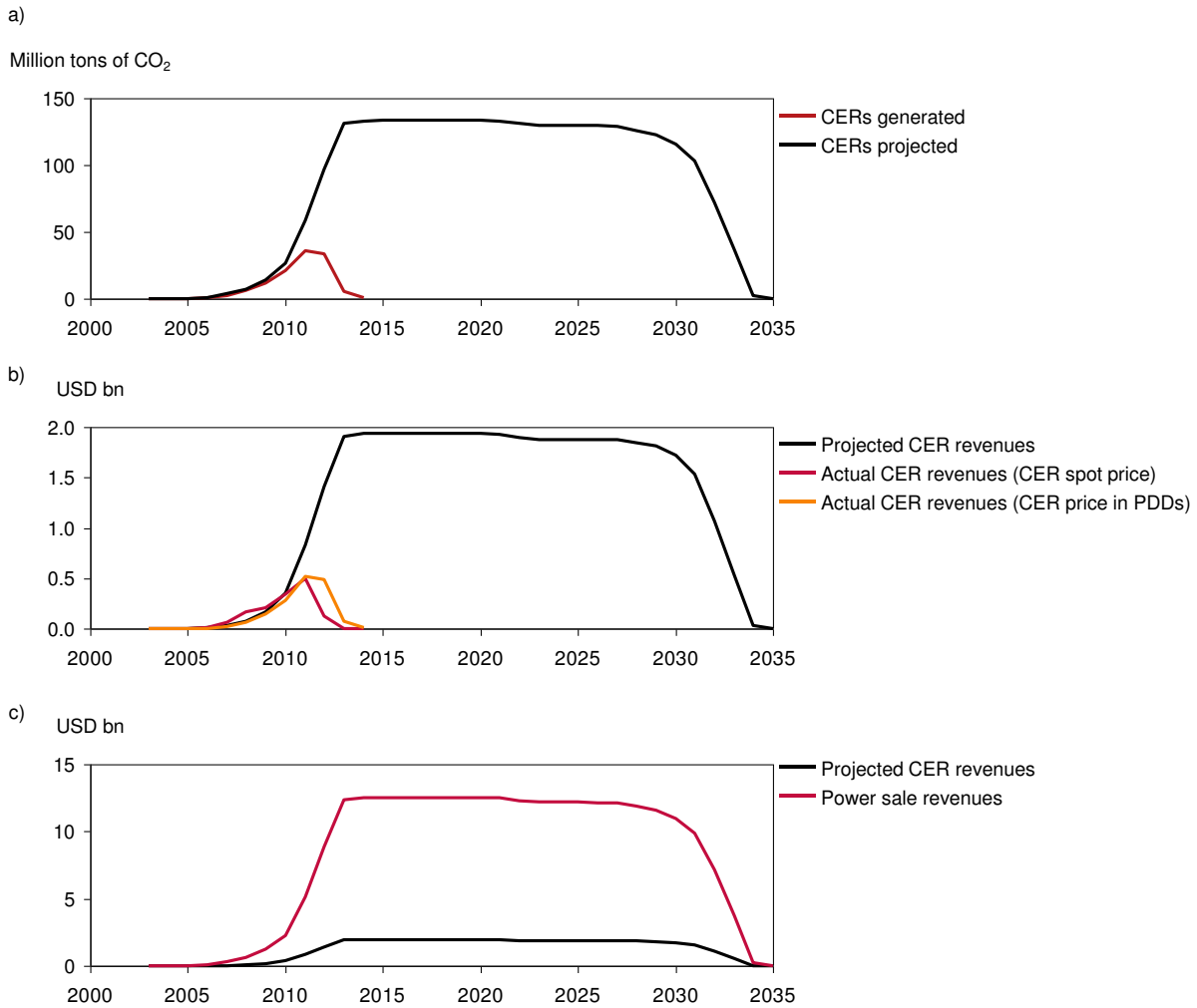
The gap between projected and actual revenues from CER sales is almost as large as in the case of emissions (see subpanel (b) of Figure 3.9). At the CER prices assumed in each project's PDD (between \$7 and \$32, with an average of \$14.6), these CERs would have yielded a revenue stream of \$1.94 billion per year and a total of \$40.09 billion until 2035. While no detailed public information is available on realized sales volumes and prices, it is clear that real revenues are falling far short of this projection. Assuming a price band established by the prices listed in the PDDs (IGES, 2015b)– which often indicate longer-term purchase contracts with buyers from industrialized countries – and the spot market price of CERs (Thomson Reuters, 2012)<sup>39</sup>, the actual *cumulative*

---

<sup>38</sup> Monitoring reports have to be filed and verified before CERs are issued, and project operators are free to choose when to report their emission reductions and in what frequency (the minimum is to report once every crediting period, i.e., once every 7 years, which means that credits can be issued for multiple previous years at once.

<sup>39</sup> We approximated CER spot prices by secondary CER prices with vintage December 2012 (for the years 2008-2012) and December 2014 (for the years 2013-2014). For the years 2003-2008 we assumed prices at 2008-levels.

revenues were somewhere in between \$1.13 billion and \$1.87 billion for 2003-2015, lower than the *annual* projected value for 2015 that was assumed by project developers when they applied for the CDM.



Data from IGES (IGES, 2015a, 2015b).

**Figure 3.9:** CER Generation and Revenue by Chinese Wind CDM Projects

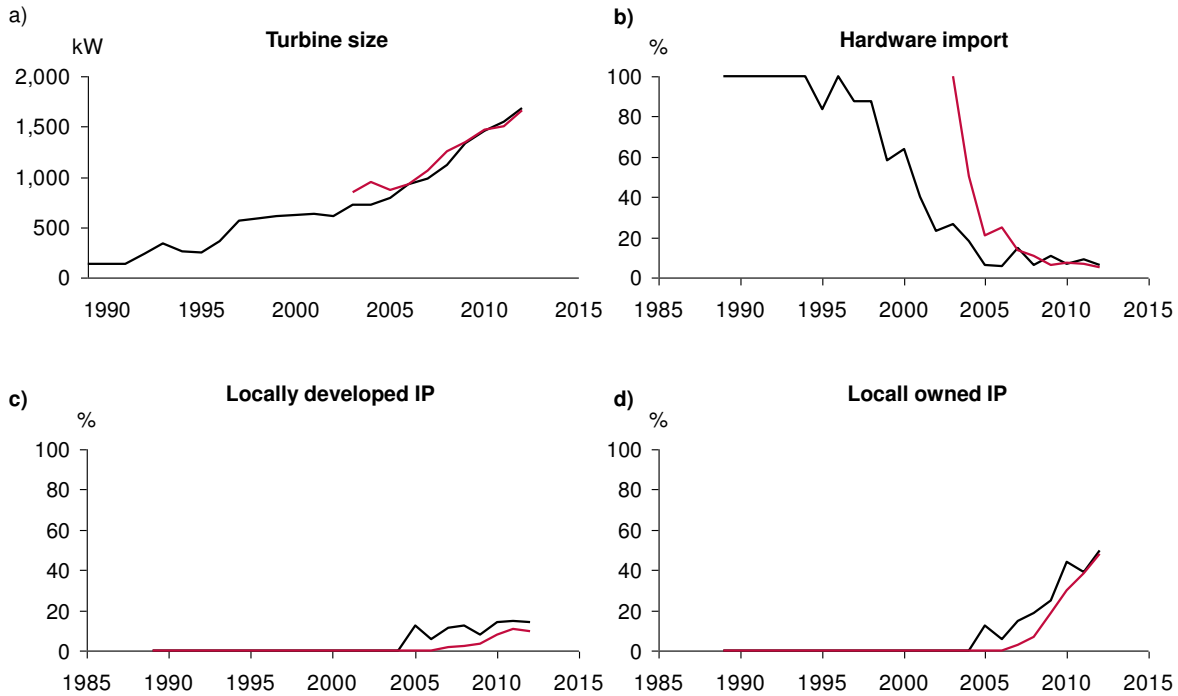
Even if projected volumes and price were to be realized in the coming decades, the CDM's contribution to project revenues is likely to be dwarfed by revenues from power generation. Subpanel (c) of Figure 3.9 shows projected CER revenues in comparison to

electricity sales, which we estimated using projected CER generation (see above) and the grid-emission factors and power tariffs listed in each project's CDM documentation. It can be seen that power sales reach approximately \$19 billion per year at their peak in 2014-2021, adding up to \$259.64 billion over 2003-2035. Even if realized at full scale, and if transaction cost for CDM monitoring, reporting and verification were zero, CER revenues would thus only contribute 13.37% of total project revenues. This is not to say that expected CER revenue was irrelevant to project planning, as typical CER prices were expected to increase project internal rate of returns by approximately 4.7 percentage points. However, CER revenues are much more uncertain than revenues from electricity sales, and in practice, the assumption that expected CER prices would persist for the lifetime of the projects has not been close to reality.

### **3.2.3.3 Technology Choice under the CDM**

In addition to the direct effect of CDM revenue on project design, there is also evidence that the prospect of CDM registration influenced technology choices by investors. Figure 3.10 compares CDM and non-CDM projects on four dimensions of technology choice: the size of turbines (in kW; subpanel (a)), whether turbines were directly imported from abroad (subpanel (b)); and whether turbines with locally developed IP (subpanel (c)) or locally owned IP (subpanel (d)) were employed. The graphs illustrate that CDM projects used – on average – larger turbines and more often foreign technology, especially in the early years of the CDM. But the differences are marginal, and seem to become less significant over time: all technology sourcing modes are present in the population of CDM and non-CDM projects from 2006 onward, when annual investments pass the billion-USD-per-year threshold. This confirms previous accounts of the influence of the CDM on technology transfer (Lema and Lema, 2013). It is

interesting to note, however, that non-CDM investors were the first to adopt locally developed and locally owned IP on a significant scale, a trend which CDM investors later followed.



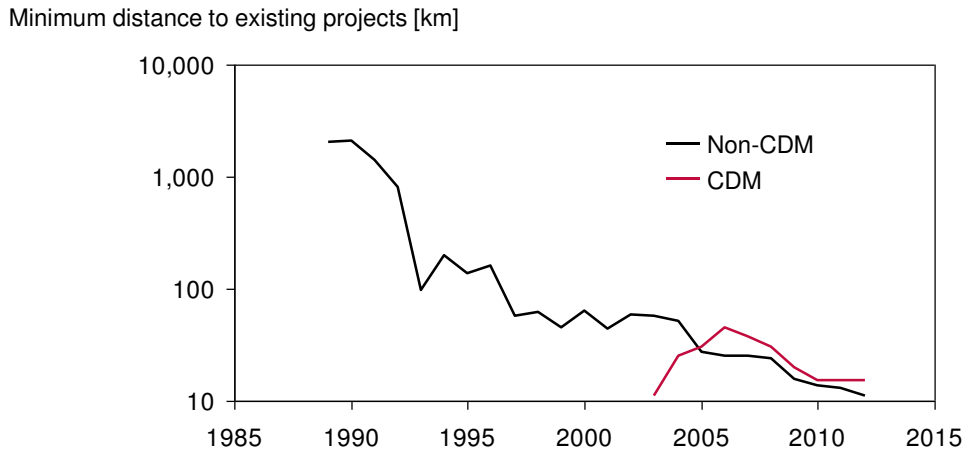
**Figure 3.10:** *Differences in Technology Choice between CDM and non-CDM Projects*

### 3.2.3.4 Geographical Diffusion of Wind Projects

Even more than turbine technology choice, siting the wind farm is the most important project decision for wind investors. But new and potentially more lucrative wind resources are often located in remote and difficult terrain and it takes time to fully assess and understand a particular location's wind resource quality. These factors increase investment risk. Further, being close to existing wind farms often means that access roads already exist or can be extended, and that installation and maintenance resources – such as cranes and trained personnel – are available at lower cost. Wind



farms in China and elsewhere are therefore often developed in clusters of several hundred or thousand MW, and investors are hesitant to develop projects in new locations. Did the CDM make such investments in previously undeveloped areas more attractive? Our results suggest that there is evidence for this effect. The average minimum distance to existing projects, shown in Figure 3.11, is lower for non-CDM projects than for CDM projects in all years from 2005 to 2012 (the period of highest investment). Weighted by installed capacity, CDM projects were on average 17.12 km away from existing wind farms, compared to 14.65 km for non-CDM projects (in 2003-2012; see Figure 3.11). But here, too, the effect appears relatively small compared to the general sector-wide trend.



Average minimum distance of new investments to existing wind projects for CDM and non-CDM projects. For both classes of projects, the average distance decreases over time as more and more resources are exploited.

**Figure 3.11:** *Average Distance of New Wind Projects to Closest Existing Projects*

## **3.2.4 Discussion**

### **3.2.4.1 Shifting Trajectories of Development**

What was the impact of the CDM on the development of the Chinese wind sector? Since all indicators of technological change in the Chinese wind sector are changing rapidly as deployment has rapidly grown, disentangling the impact of the CDM is difficult when looking only at sector-level indicators of technological change. Our analysis of micro-level techno-economic data provides a novel perspective and new insights on the development trajectories in the sector and the role of the CDM relative to domestic subsidies from the Chinese government.

Our analysis shows that about 80% of Chinese wind projects is registered with the CDM. A total of \$82 billion in capital investment was made with support from the CDM. These projects are projected to generate the equivalent of 2.8 billion tons of CO<sub>2</sub> reduction credits for the global carbon market. The CDM thus provided – and will continue to provide – additional revenue for the large majority of Chinese wind farms.

However, while capital investments under the CDM were substantial, our results illustrate that the populations of CDM and non-CDM projects followed remarkably similar trajectories in 2003-2012 on all indicators we analyzed. Investment in wind projects grew exponentially inside and outside the CDM; the average size of turbines increased steadily over time; investors increasingly relied on locally sourced technology; and investors gradually exploited the country's dispersed wind resources. The projects the elected into the CDM seem to have followed a strategy of intensifying development rather than changing the trajectory of development. It is therefore useful to think of the impact of the CDM in terms of shifts in the development trajectory: Did the CDM advance development on one or several of the indicators of technological change?

Under the UNFCCC's supposition that the CDM projects were all financially additional – and this is a major assumption we evaluate directly in Section 3.3 – the marginal capital investments attracted through the CDM pushed the capacity trajectory of the Chinese wind sector ahead by about 3.5 years. This means that due to the CDM, wind installations reached 16 GW – the level of non-CDM installations without the CDM at the end of 2012 – already in mid-2009.

How material is a shift of 3.5 years in the development of a sector? Under the one-for-one substitution dynamics of energy markets assumed under the CDM, the \$82 billion of investments in wind power under the CDM in China reduced demand for investment in fossil-fuelled alternatives. If the energy demand fulfilled by the CDM wind projects would instead have been satisfied by coal-fired power plants, the 64 GW of wind would have replaced around 15-20 GW of coal, assuming typical capacity utilization factors for coal and wind in China. This is roughly equivalent to the coal-fired power plant fleet of the United Kingdom. Further, this assumption of one-for-one substitution yields even more important potential environmental consequences, because the emissions trajectory of China would have been altered for several decades as new coal capacity would have locked-in emissions for the lifetime of the plants, typically 40 years (Davis and Socolow, 2014). This suggests that there might be a role for offsetting mechanisms in the development trajectory of emerging economies, most notably India, which will face rapidly growing energy demand in the next decades that will require short-run decisions on new energy generation capital that will have long-run environmental consequences.

Again under the assumption of full financial additionality, the CDM appears to have helped accelerate the trajectory of turbine-sizes, especially in the years 2003-2008, when CDM projects were up to 2.5 years ahead of the non-CDM projects. The CDM also

appears to have provided significant incentives to keep importing advanced technology from abroad at a time when local-content requirements and cost-competition imposed pressure to invest in less advanced domestic technology. In terms of geographical diffusion, CDM projects appear to have been more risk-taking than non-CDM projects. Taken together, the indicators on deployment, specific turbine size, local manufacturing, and turbine development point toward the promise of international offsetting mechanisms on the development of low-carbon technology sectors in emerging economies. However, this promise is only a reality if the CDM projects were financially additional, otherwise, these trends are not a reflection of the CDM as a source of finance, but rather a spurious relationship due to the differential ability of project developers who did or did not participate in the CDM. Further, it is of note that on most indicators, the technological trajectories of CDM and non-CDM projects have converged in recent years, in particular the trajectories of technology choice and technology sourcing.

### **3.2.4.2 Impact of the CDM on the Financial Bottom-Line**

The fact that CDM and non-CDM projects differ slightly on averaged characteristics hides that there is large overlap between the two populations on all indicators. This suggests that some of the projects that received CDM funding might not have been truly additional and should not have been able to pass the UNFCCC's additionality criteria (see Section 3.3.3.1), and that some projects that did not receive CDM funding would have been able to register successfully<sup>40</sup>. The fact that investment continued almost unabated after 2012 (see Figure 3.12), the last year projects were able to register for the

---

<sup>40</sup> We do not observe organizational factors, e.g., if the project owner is state-owned or privately owned, which might influence the ability to obtain the official approval required to apply for the CDM. The reasoning explaining why not all projects that appear eligible for the CDM participated in the system remains an open question.

CDM and be able to sell into the European Union's carbon market, also raises doubts about how important the CDM was for investment decisions in the Chinese wind sector.

### **3.2.4.3 Future Research**

This chapter points toward two avenues for future research. *First*, there is a need for micro-level and macro-level studies of the additionality of CDM financing. Future research should make use of micro-level wind resource data and the financial information contained in the CDM documentation to assess financial additionality of Chinese wind projects. More generally, so far there has been too much focus in the literature on the assessment of investments that did receive CDM funding and too little research on investments that did not. A better understanding of the counterfactual investment trajectories absent the CDM is needed to better the net impact of the CDM on emission trajectories of developing countries. *Second*, our findings point toward the opportunity to exploit the very rich, micro-level data available in the CDM documentation to analyze the Chinese wind sector more broadly. Tang and Popp (2014) took a first step in this direction by analyzing learning effects. But the data, which includes company information, tariffs, taxes, cost, loan conditions, as well as detailed technical information, may prove useful for a number of research questions relating to public policy and investment decisions in the world's largest wind sector and investments in renewable energy in emerging economies.

### **3.2.5 Conclusions and Policy Implications**

The Intergovernmental Panel on Climate Change (IPCC) emphasizes the need to rapidly scale-up investments in low-carbon energy technologies in emerging economies in order to mitigate dangerous climate change (IPCC, 2014). Better understanding the

influence of international climate policy on the development of China's wind sector –the largest non-hydro renewable energy sector in the world – is therefore highly relevant to inform the design of the domestic and international climate policy regime post-2020.

This section presents a novel, project-level analysis of Chinese wind investments that allows us to identify which projects received CDM funding and how they differ from projects that did not. The comparison of the aggregate and project-specific characteristics of investments under the CDM with investments realized without CDM support allows us to draw conclusions about the marginal impact of CDM funding on investment and project-design decisions.

Our four main findings are that (i) a large majority – about 80% – of Chinese wind investments between 2003 and 2012 received CDM support; (ii) CER revenues have accounted for a relatively small share of total project revenue; (iii) CDM projects were, on average, slightly more likely to invest in more advanced foreign technology and slightly more likely to exploit previously undeveloped wind resources; and (iv) the characteristics of CDM and non-CDM projects tend to converge toward the end of the observed period.

Under the assumption that all projects were financially additional, the marginal capital investments attracted through the CDM pushed the capacity trajectory of the Chinese wind sector ahead by about 3.5 years. This points toward the positive effect that international offsetting mechanisms can potentially have on the development of low-carbon technology sectors in emerging economies. However, apart from absolute capital investment, the impact of the CDM on the long-term trajectory of the sector appears to have been marginal. The CDM's impact seems to have advanced the sector's

development rather than changed its trajectory, and on most indicators, the technological trajectories of CDM and non-CDM projects converged in recent years.

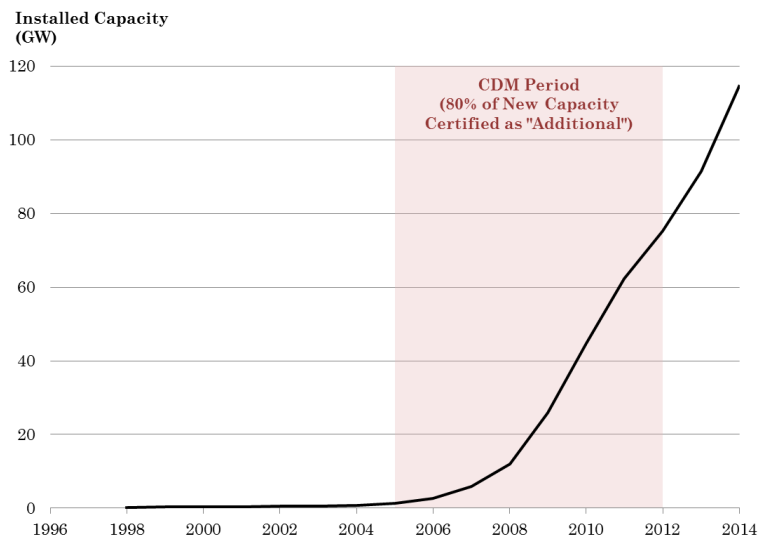
The actual role of the CDM in the development of China's wind sector thus hinges critically on the concept of financial additionality. Yet at the same time, the similarity of CDM-projects and non-CDM projects on key project indicators averaged over a large number of projects suggests that some of the projects that have received CDM funding might not have been truly additional, and that some projects that did not receive CDM funding would have been able to pass the UNFCCC's additionality criteria. This raises two key questions: First, what share of projects were truly additional? And why did those projects that did not receive CDM funding not apply for it? These questions have potentially significant implications for the adequate role of offsetting mechanism in international climate policy in general and need to be explored in further research.

### **3.3 The Additionality of Clean Development Mechanism Projects in the Chinese Wind Sector**

The environmental effectiveness of an offset scheme in practice depends centrally on whether or not the emission reduction projects financed through the CDM would have occurred in the absence of the CDM, as described in Section 3.1.1. This criterion has been formalized in the design of the CDM through the "additionality" condition. In this section, we present an evaluation of additionality in one of the largest sectors of the CDM, the Chinese wind sector. In this section, we take a variety of methodological approaches to evaluating additionality, and then use the results of our assessment to offer insight about the design of future offset policies. The context for this analysis is the Chinese wind sector, which has seen the most rapid growth of any national non-hydroelectric renewable energy sector, as discussed in Section 3.1.3. As shown in Section

3.2, the period of intense sectoral growth in 2005 – 2012 saw 80% of the projects partially financed through the CDM. Figure 3.12 puts this into perspective, showing the background of rapid wind sector growth and the period during which new projects could apply for financing through the CDM. CDM projects were required to demonstrate that their construction would not have been financially viable without the additional support of the CDM. However, when the Chinese projects were no longer eligible for CDM finance in 2013-2014, the development of the wind sector continued in what appears to be an uninterrupted trend.

In this section, we explore the additionality claims of wind energy projects during this time period utilizing the carefully matched wind project database, whose construction is described in Section 3.2.



The red box corresponds to the period 2005 – 2012 when new projects were eligible to apply for finance through the Clean Development Mechanism (CDM). Following 2012, the rapid trend in wind energy development appears uninterrupted despite the fact that 80% of projects constructed from 2005 – 2012 established that they were “additional” and were only financially viable with the support of the CDM. Projects built starting in 2003 were eligible to apply for the CDM, but there was significant uncertainty in the viability of the scheme until 2005.

**Figure 3.12:** *Cumulative Installed Wind Capacity in China from 1998 – 2014*



### **3.3.1 Introduction**

The additionality of CDM projects is of central importance to the environmental effectiveness of the CDM and the Kyoto Protocol. Understanding the additionality of CDM projects can inform the design of future offsetting schemes and the viability of future project-based schemes. In the CDM, additionality has been institutionalized in practice through a wide variety of approved methodologies for evaluating projects and relevant baselines. As described in Section 3.2, the Chinese wind sector has made up a large fraction of total CDM projects globally. As in many sectors of the CDM, a dominant approved methodology has emerged for certifying Chinese wind CDM projects.

For 98.4% of Chinese wind CDM projects, additionality was assessed with a single methodology, ACM0002, known as the “Large-scale consolidated methodology: grid-connected electricity generation from renewable sources.” Under ACM0002, renewable energy projects can establish their additionality by conducting “investment analysis,” which requires demonstrating that the CDM project in question satisfies an established “benchmark.” In the practice of CDM in the Chinese wind sector, this has led to an institutional development whereby nearly every project justifies its additionality in terms of the project finance statistic, internal rate of return (IRR). In practice, Chinese wind CDM projects have been certified as additional by the CDM Executive Board if they can demonstrate that the IRR of the project is below an established benchmark without financial support of the CDM but is above the benchmark with CDM finance. In practice, a single number has been used for all wind projects in China over the lifetime of the CDM: eight percent. Approved projects then generate credits proportional to the average GHG intensity of electricity generated on the grid that a wind project connects to.

The validity of additionality claims hinges on the accuracy of how counterfactuals are modeled. The implicit assumption in the dominant methodology used to approve Chinese wind CDM projects is that all projects with an IRR below 8% without the CDM (and above 8% with the CDM) would not be financed independent of the CDM. Further, it is assumed that by providing the additional finance needed to become attractive to investors, these projects displace the average unit of electricity on the grid. There are several conceptual issues with the assumptions embedded in this approach.

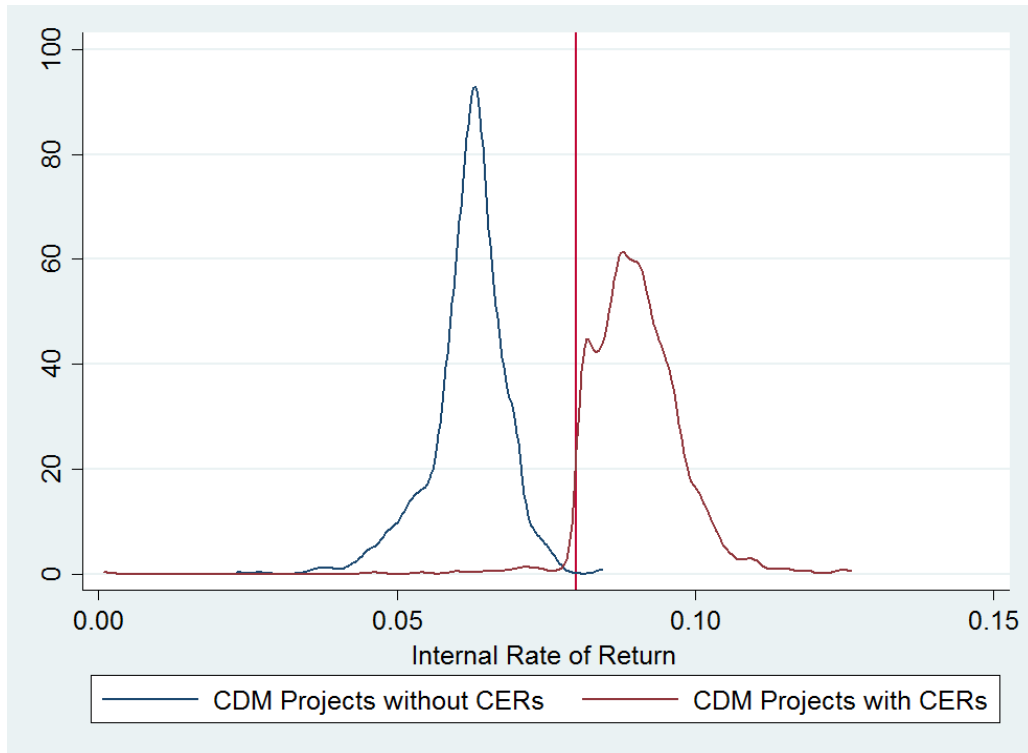
First, any additional project does not displace an average unit, but rather marginal units on both the intensive and extensive margins. On the intensive margins, nuances of electricity dispatching determine the GHG-intensity of the electricity that is actually displaced by a new investment. On the extensive margin, new investments that are not made as a result of the additional investment are displaced and the dynamics of dispatching subsequently change, which together determine the emission additionality of a project.

Second, financial benchmarks are more nuanced than the one-number summary provided by IRR. While the IRR of a project does capture much of a project's relevant financial information, there are many known limitations of IRR, such as scale-independence and blindness to opportunity cost, which can lead projects with high IRR to not be financed in practice.

Third, any single benchmark that could be used in practice will miss important nuances and heterogeneity in the actual financial hurdle rates of projects in practice. The investment environment of a project depends on temporal and spatial dynamics, thus a single benchmark can become both outdated and irrelevant if not closely updated and tailored for each project based on empirical observations of financial conditions.

Finally, the benchmark that has been institutionalized over the duration of Chinese CDM wind activity is both outdated and largely irrelevant to the projects it is used to inform. The 8% benchmark for IRR is derived from a 2002 report of the State Power Corporation of China titled, “Interim Rules on Economic Assessment of Electrical Engineering Retrofit Projects,” which posited (without empirical justification) an acceptable IRR to be used during the State Power Corporation’s breakup as a vertically integrated monopoly (SPC, 2002). For a thorough discussion of the 8% benchmark, see He and Morse (2013, 2010), who state that this benchmark is “taken from an interim decree of a defunct vertically integrated monopoly [and] cannot be considered legitimate for CDM additionality determination in the current Chinese market.”

The effect of the 8% benchmark rule is reflected in the distribution of self-reported IRR’s in the PDDs. Figure 3.13 displays the distribution of IRR’s reported in PDDs, as calculated for scenarios with and without revenue from CERs. The figure shows that without CER revenue, self-reported IRRs are highly concentrated in the range of 5 – 7%, and with CER revenue, IRRs are above 8%. The figure suggests that the 8% threshold is binding, in the sense that with CER revenue, reported IRR’s are concentrated just above the threshold. The apparent censoring at the 8% benchmark is suggestive of either the CDM process effectively screening for the projects that are truly additional or strategic self-reporting of project financials to satisfy the criteria established by the 8% benchmark.



The left (blue) distribution shows the reported IRR for projects in the absence of revenue from the generation of certified emission reductions (CERs) and the right (red) distribution shows the reported IRR when CER revenue is included. The horizontal red line is placed at the 8% IRR benchmark that was used to determine CDM eligibility for over 98% of Chinese wind CDM projects.

**Figure 3.13:** *Self-Reported Internal Rates of Return for Chinese Wind CDM Projects*

The conceptual issues with how additionality has been assessed in practice for Chinese wind CDM projects are challenging to overcome. As discussed in Section 3.3.4, this institutional structure has led to perverse incentives and strategic manipulation of regulations. In the Chinese wind sector context, we are able to apply a significantly more rigorous approach to modelling counterfactual scenarios to inform an assessment of additionality. As discussed in Section 3.2.4, 20% of Chinese wind projects financed during the time period of the CDM did not receive support from the CDM. By definition, these projects are non-additional – they were built without the support of the CDM. Leveraging data on these projects can enable an empirically grounded model of counterfactual wind development that has significantly greater analytical rigor relative

to the dominant institutionalized approach. With data on both CDM projects and non-CDM projects, as described in Sections 3.2.2 and 3.3.2, we are able to create several models of counterfactual wind project characteristics that serve as meaningful comparisons in additionality assessments. These alternative models rely on less and more restrictive assumptions, as described in Section 3.3.2. This section presents the findings of applying these counterfactual models to a comprehensive dataset of Chinese wind projects. Sections 3.3.4 and 3.3.5 draw out conclusions from these results and apply these findings to institutional reform of future offsetting schemes.

### **3.3.2 Data**

Section 3.2.2.1 describes the two main databases of wind energy projects and characteristics that we use in this work: the Huaxia database and the CDM database. For the analysis of additionality, we integrate this data with additional matched data on wind resources. Assessing additionality requires careful consideration of potential energy generation, which depends on project characteristics (e.g. turbine choice) and site characteristics. To capture relevant site characteristics, we incorporate data from three databases of wind resources: a wind resource assessment conducted by the Harvard China Project (McElroy et al., 2009), a commercial product developed by the company 3TIER, called the Wind Time Series and Prospecting Tool (3TIER, 2015), and a dataset of near-ground wind speeds reported by NASA (NASA, 2005).

#### **3.3.2.1 Wind resource estimation with cross-validation**

The three wind resource databases we utilize (henceforth “3TIER,” “Harvard China Project,” and “NASA”) partially overlap and share much of their underlying data and methods, yet differ in several key aspects. The datasets rely on satellite observations of

mid-atmosphere wind patterns, but differ in how this data is down-scaled to a geographic resolution relevant for assessing wind resources available to a specific wind farm. The three datasets use different approaches to down-scaling and rely on different topographical datasets, which is a crucial determinant of a wind farm's effective wind resource. Topographical data determines the surface roughness in a particular region and determines the turbulence that affects the effective wind resource that can be extracted by a turbine. The three datasets provide different measures of wind resources and we use standard engineering formulae to map estimated wind speeds to capacity factors.

While correlated, these highly related data sources offer alternate projections for wind resources that capture different information. Rather than building a new physical model of wind resource, we utilize these three models in a cross-validation model selection approach, using observed generation as the calibrating baseline. We collected generation data for 492 CDM projects from 2005 – 2012 from the CDM's monitoring and verification reports. For these projects, we collect 1,045 observations of annual energy generation fed into electric grids, as verified by third party auditors approved by the UNFCCC. This observed generation data allows us to assimilate the wind resource data, utilizing additional observed controls, to build a predictive model of energy generation that can then be used to forecast generation for the remaining projects, including the non-CDM projects.

Our cross-validation approach to assimilation model selection uses an 80-20 leave-out rule, where we randomly take a subsample of 80% of the 1,045 observations without replacement, called the "training" set, fit several regression models using the various wind speed measures and control variables to predict observed generation, and repeat

this procedure for 1,000 re-sampled training sets. Regression models are then evaluated based on the root mean squared error (RMSE) in predicting the generation of the held out 20% of the data not included in the training set<sup>41</sup>. Cross-validation avoids over-fitting by penalizing predictions driven by a small number of training set observations that do not represent the relationships in the held-out set.

The general cross-validation regression is specified in Equation (3.1), where  $CF_{3TIER_i}$  is the capacity factor estimated with the 3TIER data at location  $i$ ,  $CF_{HCP_i}$  is the capacity factor estimated with the Harvard China Project data at location  $i$ , and  $CF_{NASA_i}$  is the capacity factor estimated with the NASA data at location  $i$ . The alternative models we explore vary the functional forms of the  $f, g, h$ , and  $j$  functions. The functional forms we explore for these functions are tabulated in Table 3.3.

$$CF_{obs_{i,t}} = f(CF_{3TIER_i}) + g(CF_{HCP_i}) + h(CF_{NASA_i}) + j(controls_i). \quad (3.1)$$

**Table 3.3:** Cross Validation Functional Forms

$f, g, h(CF_i)$		$j(controls_i)$	
Model	Specification	Model	Specification
I	$\beta_1 CF_i$	I	$size_i + maker_i + year_i$
II	$\beta_1 CF_i + \beta_2 CF_i^2$	II	$size_i + maker_i + year_i + province_i$
		III	$\beta_1 size_i + \beta_2 size_i^2 + maker_i + year_i + province_i$

Model numbers and specifications for the  $f, g, h$ , and  $j$  functions of the regression models considered in cross-validation. In the  $j$  functions,  $size_i$  is the size of the turbine (in MW),  $maker_i$  is the manufacturer of the turbine,  $province_i$  is the province the turbine is installed in, and  $year_i$  is the commissioning year of the project. All three models use fixed effects for turbine size, maker, and commissioning year. Models I and II of the  $j$  function include fixed effects for turbine size (there are 13 distinct turbine sizes in the data set), whereas Model III utilizes a quadratic function of turbine size. Models II and III of the  $j$  function also include province fixed effects.

---

<sup>41</sup> Regression models could also be evaluated in this context using the  $R^2$  statistic. RMSE is preferable to  $R^2$  in this context because we are interested in the magnitude of deviations between predictions and observed values, rather than relative deviations. Nevertheless, our results are qualitatively robust to using  $R^2$  as the evaluation statistic rather than RMSE, although there are a few small flips in the rank-order of models – none of which flip for meaningfully large differences in  $R^2$ .

We assess 29 models of the form described in Equation (3.1) that combine the functional forms described in Table 3.3. The full set of results from cross validation is shown in Table 3.4. The dominant combination of functional forms for the  $f, g$  and  $h$  functions is Model II for equations  $f$  and  $g$ . Interestingly, the preferred model excludes equation  $h$  (effectively dropping the NASA data). Across the three models for equation  $j$ , the dominant model in terms of mean RMSE is Model III. Together, our preferred assimilation model is the one with the lowest mean RMSE from the results of 1,000 cross-validation simulations, given by Equation (3.2):

$$\begin{aligned} \widehat{CF_{obs}_i} = & \hat{\beta}_0 + \hat{\beta}_1 CF\_3TIER_i + \hat{\beta}_2 CF\_3TIER_i^2 + \hat{\beta}_3 CF\_HCP_i + \hat{\beta}_4 CF\_HCP_i^2 + \hat{\beta}_5 size_i \\ & + \hat{\beta}_6 size_i^2 + maker_i + year_i + province_i \end{aligned} \quad (3.2)$$

Notably, Equation (3.2) describes the model with the highest predictive power, but excludes all information from one of our three wind speed data sources, the NASA dataset. The finding that one of the three datasets does not improve predictive power for out-of-sample locations highlights the utility of the cross-validation approach. As an alternative, including all information in a predictive model would guarantee higher in-sample predictive power, but, as these results indicate, would lead to over-fitting and less accurate out-of-sample prediction.

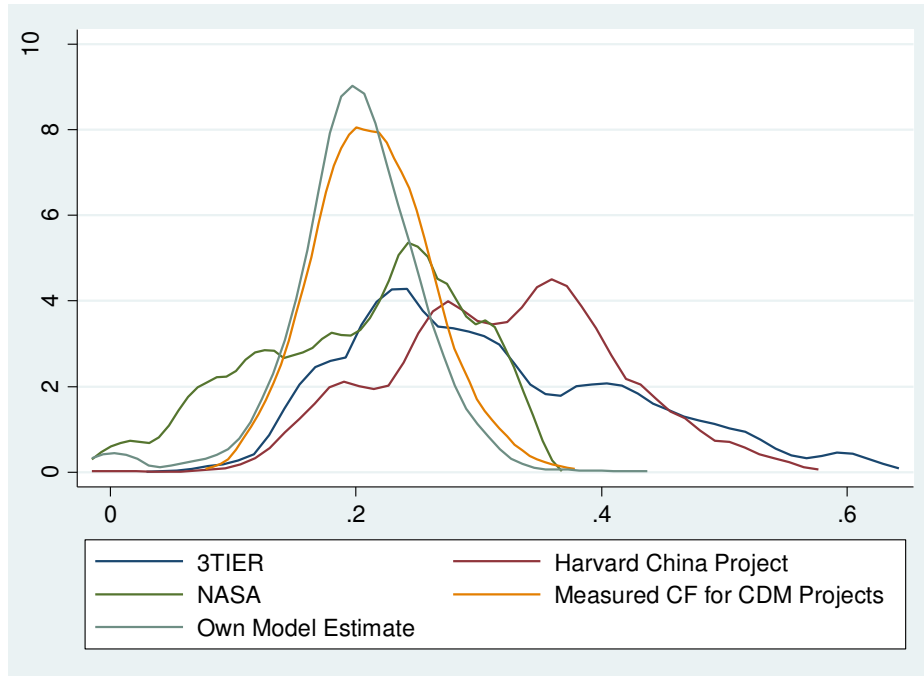
Although not the focus of our work, this calibration model approach can also be used to assess the accuracy of wind resource models and may be a useful direction for future research in developing wind resource models that utilize observed generation data at similar sites in wind development prospecting and academic research on wind generation potential.



**Table 3.4:** *Cross Validation Model Evaluation*

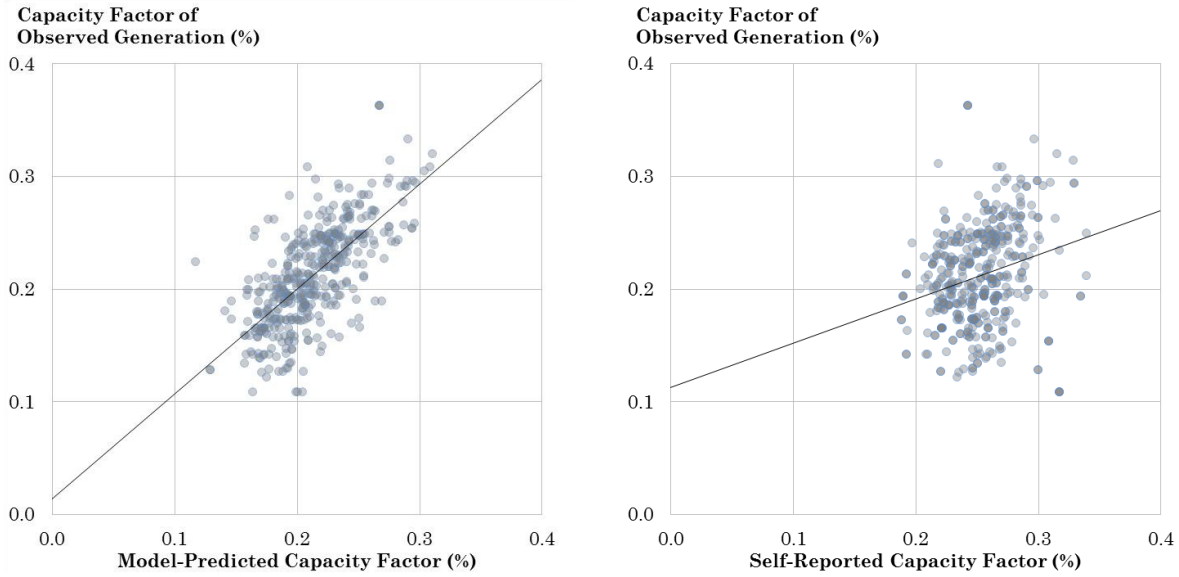
$f$	$g$	$h$	$j$	Mean RMSE	Median RMSE
-	-	-	III	0.04113	0.04116
I	-	-	III	0.03919	0.03921
-	I	-	III	0.04016	0.04012
-	-	I	III	0.04060	0.04064
I	I	-	III	0.03926	0.03932
I	-	I	III	0.03922	0.03927
-	I	I	III	0.04002	0.04003
I	I	I	III	0.03928	0.03932
II	II	II	III	0.03891	0.03895
II	-	-	III	0.03910	0.03913
-	II	-	III	0.03954	0.03958
-	-	II	III	0.04058	0.04060
II	II	-	III	0.03887	0.03891
II	-	II	III	0.03912	0.03916
-	II	II	III	0.03949	0.03951
II	II	II	II	0.03977	0.03962
II	-	-	II	0.03982	0.03968
-	II	-	II	0.04045	0.04030
-	-	II	II	0.04134	0.04129
II	II	-	II	0.03976	0.03961
II	-	II	II	0.03982	0.03965
-	II	II	II	0.04038	0.04022
II	II	II	I	0.04109	0.04097
II	-	-	I	0.04132	0.04130
-	II	-	I	0.04405	0.04409
-	-	II	I	0.04414	0.04412
II	II	-	I	0.04104	0.04093
II	-	II	I	0.04120	0.04114
-	II	II	I	0.04395	0.04402
Minimum RMSE				0.03887	0.03891

The projections from the three wind resource data sets, measured capacity factors, and projected capacity factors from the model displayed in Equation (3.2) are shown in Figure 3.14. Figure 3.15 compares the model-predicted capacity factor and self-reported capacity factor in PDDs to the observed capacity factor from generation data.



Capacity factor estimates from three data sources (3TIER, the Harvard China Project, and NASA), measured capacity factors for a subset of these projects, and the fitted capacity factors from the preferred model described in Section 3.3.2.1. By calibrating with the observed capacity factors, our model estimates for all projects closely approximates the mean, variation, and distribution of observed capacity factors.

**Figure 3.14:** *Capacity Factor Estimates and Observations*



Comparison of observed capacity factors (yearly averages for all CDM projects that have at least one monitoring report with generation measures) with model-predicted capacity factors (left panel) and with self-reported capacity factor estimates in project planning documents (right panel). The model-predicted capacity factor, which is calibrated to the observed generation with available wind resource estimates, captures 48% of the variation in observed wind speeds, whereas the self-reported capacity factors capture 6.1%. For comparison, the natural year-to-year variation in capacity factors is such that previous annual generation explains between 37 – 67% of future annual generation.

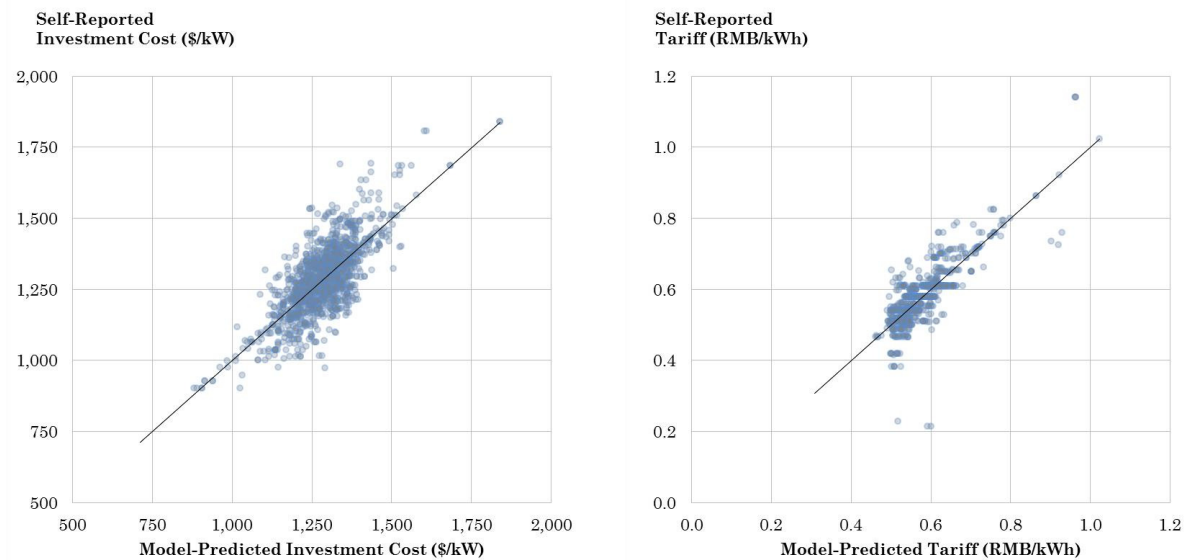
**Figure 3.15:** *Comparison of Capacity Factor Estimates and Observations*

### 3.3.2.2 Interpolation of Variables Using Project Design Documents

In addition to wind speed, two other financial variables are available in the CDM project design documents but not available in our database on non-CDM projects: investment cost and tariffs (the electricity price guaranteed to project developers). Rather than using fixed values for these variables, we estimate a parametric relationship between these variables and variables we observe for both CDM and non-CDM projects. We adopt a similar cross-validation approach to the one described in Section 3.3.2.1 to allow us to test a large number of functional form relationships, while remaining agnostic to the underlying structural relationships between these parameters. For example, tariffs are typically set by provincial governments and are typically

negotiated based on observable project characteristics and developer experience. We observe these relationships and can flexibly estimate a parametric model to extrapolate expected tariffs for the non-CDM projects that do not report tariffs.

We examine 36 models with various functional forms and variable-selection rules, and repeat 1,000 cross-validation models with an 80-20 leave-out rule. We select the model with the lowest average RMSE. The relationship between investment cost and tariffs projected in the cross-validation model and observed for the CDM projects is shown in Figure 3.16. For tariffs, of the 36 models explored, the model with the lowest RMSE based on the cross-validation exercise includes the commissioning year of the project, the commissioning year squared, the turbine size, the maker of the turbine, the developer, the province, the total wind farm size, the 3TIER capacity factor, and the cumulative installed capacity experience of both the manufacturer and farm owner. For tariffs. For investment cost, the model with the lowest RMSE has significant fewer observed variables: commissioning year, turbine size, turbine maker, project developer, and province.



Comparison of observed investments costs (left panel) and tariffs (right panel) to model-predictions.

**Figure 3.16:** *Modeled Investment Cost and Tariff versus Observations*

### 3.3.3 Methods and Results

The richness of the project-level data we utilize in this chapter enables us to implement several models that assess additionality. Fundamentally, because additionality relies on estimating a counterfactual, evaluation of additionality claims will be assumption-dependent. Thus, leveraging the richness of our data, we apply two types of models that vary in the restrictiveness of their assumptions. In both approaches, we rely on the merging of CDM and non-CDM datasets, described in Section 3.2, with the assumption that non-CDM projects are not additional.

#### 3.3.3.1 Exact Matching

Projects that were constructed without financial support are by definition non-additional. These projects, which are theoretically lower on the supply curve, are profitable (or otherwise viable) without additional financial support, as revealed by observed decisions. A concern in the context of the Chinese wind sector may be that

projects may be constructed that are actually not privately profitable but supported through government decisions (such as state owned enterprise decisions), and are therefore not below the market-clearing price that determines profitability. This raises a fundamental concern with additionality assessment reflected in the debate on E+/E- policies, namely the difficulty of incorporating domestic actions that affect profitability of projects in evaluation of additionality. This concern is beyond the scope of this chapter, and instead, we make the assumption that projects constructed without the support of CDM revenue should define a baseline for activities that are feasible without additional financial support. In a sense, this assumes that the posited government actions that enable these projects to be built are limitless and scale-independent (i.e. that all projects with the characteristics of the projects built without CDM support would receive the same government support and would therefore also be feasible).

With this framing, the non-CDM projects can be used to define an empirical “moving baseline” that defines the characteristics of feasible, non-additional projects. The characteristics of projects relevant for defining this moving baseline should be the most important determinants of profitability. Four of the most central parameters of wind farm project profitability are a project’s province, its commissioning date, its turbine size, and the quality of its site’s wind resources. Of the approximately 1,500 CDM projects, we find that 50.3% of the total installed CDM capacity was constructed in the same province, within 6 months, with the same turbine size, and within 1 percentage point in projected capacity factor (equivalent to approximately 25% of one standard deviation) as a project built without CDM finance. Further, of this 50.3%, 58.4% of that capacity (29.3% of the CDM total) also matches on the same turbine maker (in most cases, this implies not only a turbine of the same size but also of the same model).

These results imply that based on the most important characteristics of a wind farm that determine profitability, for over half of the wind capacity financed through the CDM, there are no meaningful differences between the characteristics of those projects and projects constructed that are by definition non-additional. Taking the non-CDM projects as a moving baseline that demonstrates what is feasible in the absence of CDM, over half of the CDM projects fail to reveal meaningful observable differences.

### **3.3.3.2 Modeling Project Financials**

The IRR of a wind farm can be assessed with a classic project finance model (Deutch and Lester, 2004). We implement a project finance model that incorporates turbine characteristics (turbine size and number of turbines), site characteristics (represented by the single statistic, capacity factor), financial characteristics of the project (electricity price/tariff, investment cost – including turbine and land costs, depreciable lifetime, minimum book value, and operations and maintenance costs), general financial parameters (share of equity, loan tenor, insurance premium, interest rate, and tax rate), and carbon-market specific parameters (average GHG-intensity of the grid and the projected market price of CERs). These parameters are related in a financial model to calculate a project's IRR, as described in Appendix C.1.

Our approach to calculating IRR standardizes the structure of project finance calculations. While CDM projects submit project design documents that provide detailed spreadsheet models of project finance to show compliance with the 8% IRR investment benchmark, general financial parameters and model structure are not standardized. While each assumption in a project design document must be referenced, there appears to be little enforcement to ensure that cited references are credible or reflective of reality.

In addition to imposing a standardized project finance model and a fixed set of general financial parameters, we also utilize our calibrated wind resource estimates, investment costs, and tariffs. Methodologies for prospectively estimating wind resources vary widely and are difficult for third parties to verify, due to variability over short and long time scales and over fine geographic resolutions. Because our goal is a consistent method for estimating wind resources at all project sites, our ability to impose a standardized model of wind resources is a crucial input enabling the comparison of IRRs for CDM and non-CDM projects. Available wind resource is one of the most important determinants of IRR; therefore, inconsistent reporting of wind resources can induce large deviations in calculated IRRs for otherwise similar projects. In effect, our modeling of project financials allows us to “re-audit” the financials of CDM projects and apply a consistent modeling approach to the non-CDM projects. With this consistent approach, we can make a direct comparison to the distribution of IRR’s for CDM projects and for the non-CDM projects, which are by definition non-additional.

The assumption at the crux of this comparison is that our estimation of wind speed, investment cost, and tariffs is consistent and unbiased for the non-CDM projects, and that our financial model captures all relevant differences that determine profitability between the two groups of projects. Under this assumption, the distribution of IRR for non-CDM projects represents the counterfactual profitability of projects that are in theory below the benchmark determining additionality. To be additional, the CDM projects would need to have a statistically lower IRR than the counterfactual established by the non-CDM projects.



Figure 3.17 displays the modeled distributions of IRR for CDM projects and non-CDM projects<sup>42</sup>. CER revenue is not included in these figures, so the calculated IRRs represents an estimate of project financials that is directly comparable. The IRR we calculate based on observable project characteristics is a reasonable proxy for the marginal cost of supply that informs the schematic supply-demand shown in Figure 3.1. To be additional, CDM projects would need to have a monotonically lower IRR relative to non-CDM projects. On average, this is not the case, as the average IRR we calculate for CDM projects is 11.15% whereas the average IRR for non-CDM projects is *lower*, at 10.59%. This average difference of 0.56 percentage points is statistically significant (p-value = 0.014), providing evidence against the supposition that CDM and non-CDM projects have the same average IRR. The more relevant comparison in this context is a one-sided t-test, against the null hypothesis that CDM projects have a lower average IRR. The p-value from this test is 0.0072, suggesting strong evidence against the claim that CDM projects have a lower IRR than non-CDM projects.

This result is confirmed by conditioning on project year. Controlling for the commissioning year of the wind farm with yearly fixed effects, the conditional difference in IRR is 0.009 percentage points *higher* for CDM projects relative to non-CDM projects with a p-value less than 0.001, suggesting that rather than being less profitable, CDM projects were actually more profitable than non-CDM projects. A similar result is obtained with yearly interactions, which shows that there is no statistically significant difference in the IRR of projects that received the CDM and those that did not for

---

<sup>42</sup> 97 projects (54 non-CDM and 43 CDM) have a calculated IRR less than zero. These projects are dropped from the analysis in this section. Possible explanations for these anomalous results are poor projections of wind resources or other financials or misreported data (this appears to explain several of these). It is also possible that the calculation of negative IRR for these projects actually reflects reality and a loss by the project investors.

projects commissioned in 2003 – 2010. In 2011 and 2012, we find that again, the CDM projects had a statistically higher IRR than non-CDM projects.

Table 3.5 displays annual coefficients for the difference in IRR between CDM and non-CDM projects from this interaction regression.

**Table 3.5:** *Annual Difference in IRR for CDM and Non-CDM Projects*

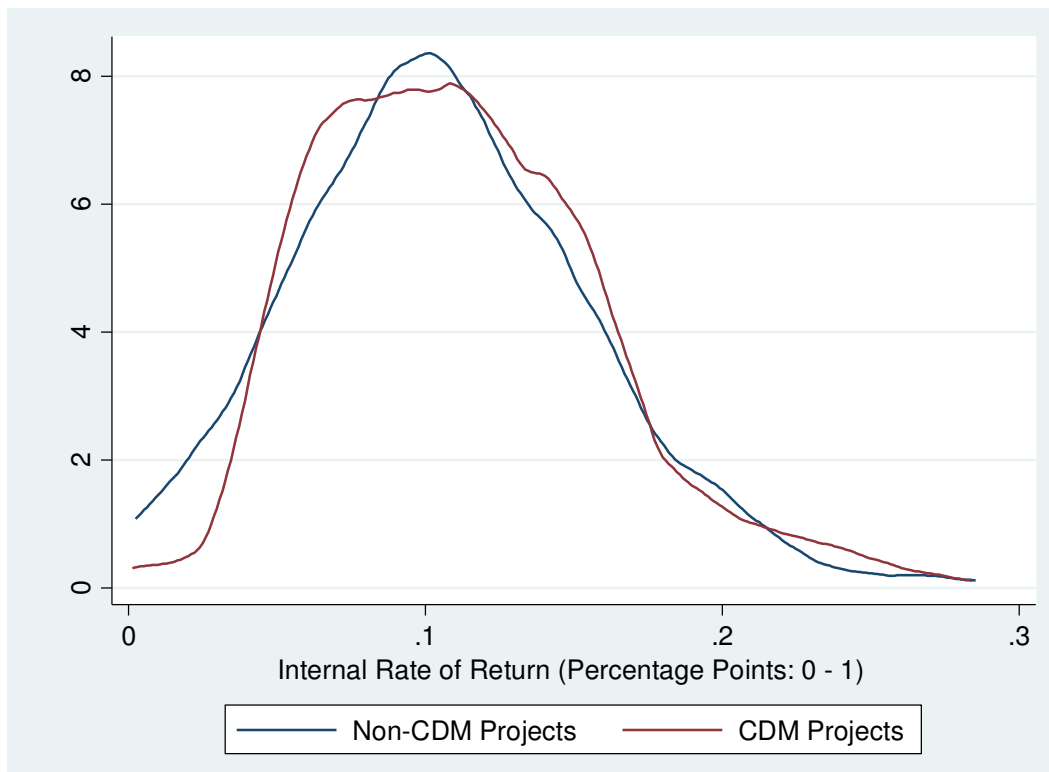
Year	Difference in IRR for CDM – non-CDM
2003	0.009 (0.069)
2004	-0.047 (0.041)
2005	0.017 (0.028)
2006	-0.001 (0.022)
2007	0.020 (0.016)
2008	-0.006 (0.010)
2009	-0.002 (0.007)
2010	-0.000 (0.006)
2011	0.021 *** (0.004)
2012	0.010 * (0.005)

\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

The table displays the annual difference in internal rate of return (IRR) for CDM and non-CDM projects (CDM IRR less non-CDM IRR given in percentage point units scaled from 0 to 1). Results are from a panel regression with project commissioning year fixed effects and individual interactions of the CDM dummy with the year fixed effects, weighting observations by the total capacity of the projects. The table presents the coefficients on the interaction terms and standard errors. Only the coefficients on receiving the CDM in 2011 and 2012 are statistically significant at the 5% level, but the signs are the opposite of what would be anticipated if the projects were additional: CDM projects in 2011 and 2012 were more financially viable without CDM financing than the non-CDM projects.

The difference in means (and conditional means) shown above is not a full assessment of additionality. The condition of additionality implies that there should be

little to no overlap in the *distribution* of marginal costs of CDM and non-CDM projects. A Kolmogorov-Smirnov test for the difference in distributions suggests that the distributions of IRR for CDM and non-CDM projects do have statistically meaningful differences with a p-value of 0.024. However, a graphical inspection of the two distributions, shown in Figure 3.17, suggests that the statistical differences are likely due to small differences in the higher moments of the distributions (i.e. the non-CDM projects have a higher variance). The distributions remain very similar after residualizing out time trends (with yearly fixed effects).



Distribution of calculated internal rate of return (IRR) for projects that received CDM financing (red) and those that did not (blue).

**Figure 3.17:** *Distribution of Calculated IRR for CDM and Non-CDM Projects*

### 3.3.4 Discussion

The Clean Development Mechanism was designed to complement the Kyoto Protocol by offering an alternative set of emission abatement opportunities that would achieve the ultimate objective of the UNFCCC (to prevent dangerous anthropogenic climate change) while also promoting sustainable development. In achieving the objective of the UNFCCC, the CDM was designed as a project-based system, with the objective of certifying every individual project on its individual environmental performance. It was a concession from the outset that requiring each project to assess its own environmental performance (by demonstrating additionality) would raise transactions costs. Nevertheless, because the CDM was to provide an alternative pool of mitigation opportunities, raising the marginal cost of CDM credits could still only improve the cost effectiveness of the Kyoto Protocol relative to not having any offset market at all.

Our results demonstrate that the environmental integrity of one of the largest national sectors of CDM activity was compromised. We find evidence that in aggregate, the wind energy projects developed in China under the CDM were not additional. While 80% of Chinese wind projects during the CDM period were supported by the CDM, 20% were not, and those 20% resemble the CDM projects in the key observable ways that determine their development costs. This result suggests that the key advantage of a project-based offset scheme, certifiable environmental performance, was not achievable in this crucial offset sector. Despite the high reporting requirement to demonstrate their additionality, essential parameters that determine the financial feasibility of projects were not successfully verified due to auditors' lack of necessary information. With a state-of-the-art wind resource dataset, reliable data on non-CDM projects, and ex-post data on CDM project performance, we bring additional information to bear that was not

available to auditors, which allows us to systematically assess additionality claims. While institutional reform to CDM rules and methodologies may be able to improve the ability of auditors to assess the additionality of projects with only data that is available ex-ante, these reforms are likely to also further increase transactions costs with no guarantee of improved environmental performance.

### **3.3.5 Conclusions and Policy Implications**

The results presented in this section highlight the immense challenges facing offsets certification. Despite the complexity of methodologies to assess additionality, ex-post assessment of CDM additionality in the Chinese wind sector suggests there may be insurmountable barriers to improving the environmental performance of project-based renewable energy offsets. In the Chinese wind sector, the dominant methodology for assessing additionality has relied on benchmarking. To be certified as additional, projects must provide a detailed financial accounting to demonstrate that their IRR would be below 8% without CDM financing but above 8% with this additional support. This suggests two fundamental issues.

First, the 8% number benchmark is not grounded in the empirical reality of current Chinese energy project finance. The 8% number is static and reliant on an ad-hoc approach. A conceptual improvement would be to update the benchmark based on empirical observations. For example, the 20% of the Chinese wind energy projects that did not receive the CDM could define a frontier of feasibility that CDM projects would have to argue they do not meet. This would be analogous to how regulations are set in the United States under the best available control technology regime.

The second issue that this form of benchmarking raises is that project developers were able to manipulate their reporting to satisfy institutional requirements that did not

reflect their project's actual additionality. Several of the parameters that have a large influence on a project's finances cannot be credibly audited by a third party. Crucially, auditors do not have a reliable third-party source for wind resource estimates. The difference in a project's potential revenue if sited in a 75<sup>th</sup> percentile location (with an average capacity factor of 23.4%) and a 25<sup>th</sup> percentile location (with an average capacity factor of 19.0%) under our wind resource estimate is 6.3 percentage points<sup>43</sup>, equivalent to 1.3 standard deviations in the distribution of IRR. Auditors are fully reliant on project developers to self-report their wind resource potential, and due to the sensitivity of IRR calculations to this parameter, the IRR that project developers reported, and by which the sole criterion for additionality is evaluated, is meaningless. This is particularly striking when it is noted that the correlation between the self-reported capacity factors (a one-number summary of the effective wind resource) and the observed capacity in the monitoring reports is 0.29 (our model-predicted capacity factor has a correlation of 0.70).

A path forward for offset markets is to consider sectoral based mechanisms. Under a sectoral mechanism, additionality criteria are established for an entire national sector. All projects in the sector are eligible for offset credit finance but at a discounted rate, to represent the aggregate difference in the sectoral baseline and the observed deployment. Sectoral mechanisms offer a possibility to dramatically lower transactions costs by removing the requirement that each individual project establish its own additionality. Instead, all projects, even those that are not additional, are eligible for offset credits, but the overall number of credits corresponds to a fraction of the emissions that a fully additional project would have offset. Sectoral mechanisms are attractive because they

---

<sup>43</sup> If all projects were located at the 75<sup>th</sup> percentile of wind resource quality, the IRR would be 13.82% in our model, but would be 7.56% if located at the 25<sup>th</sup> percentile.

lower transactions costs, but they face opposition because they do not directly assess additionality. However, in the case of the Chinese wind sector, the CDM's project-based system did not achieve its additionality objective. Fundamentally, wind energy projects may be too burdensome for auditors to evaluate at an individual project level, so sectoral mechanisms may be uniquely useful in this sector.

Sectoral mechanisms have an additional advantage in achieving the objectives of the CDM. The CDM sought to not only facilitate the achievement of the UNFCCC objective, but to also facilitate sustainable development. This suggests taking a longer, more dynamic perspective that includes considering technological change. It has been long recognized that dimensions of sustainable development that are affected by technological change, require a systems perspective, where spillovers from innovation and deployment are maximized. A project-based offset system draws the boundaries for assessment at an individual project, but spillovers fundamentally accrue across projects (and even across sectors). A sectoral mechanism would be better able to facilitate the sustainable development objective of the CDM by creating sector-wide incentives for carbon-mitigating technologies, thus encouraging a greater rate of spillover creation.

A sectoral mechanism would by no means guarantee improved environmental performance, and in fact could lead to even lower environmental performance in sectors where the project-based CDM model is performing well. However, the CDM in the Chinese wind sector is far from the efficient frontier of offset schemes depicted in Figure 3.2. A sectoral scheme would guarantee lower transactions costs, and experimentation with such a scheme could reveal that it would improve environmental performance relative to the troubling findings we've discussed in the project-based offsets in the Chinese wind sector.

### **3.4 Acknowledgments**

We would like to thank Laura Diaz-Anadon, Joe Aldy, Bill Clark, Tian Tang, Xi Lu, Daisuke Hayashi, Junling Huang for their advice at various stages of this project. We would also thank seminar participants of the Harvard Environmental Economics Lunch, and the Harvard Energy Technology Innovation Policy and Consortium for Energy Policy Seminar. We thank Charlie Wise for assistance with data management and Karl Aspelund for excellent research assistance. We thank the Harvard Environmental Economics Program, the Harvard Sustainability Science Program and the Harvard Science, Technology, and Public Policy Program for financial support. All errors are our own.



# Bibliography

- 3TIER, 2015. Wind Time Series and Prospecting Tools. 3TIER by Vaisala.
- 113th Congress, 2014. H.R. 5120 Department of Energy Laboratory Modernization and Technology Transfer Act of 2014.
- Abadie, A., Imbens, G., 2006. Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica* 74, 235–267.
- Abdulla, A., Azevedo, I.L., Morgan, M.G., 2013. Expert assessments of the cost of light water small modular reactors. *Proc. Natl. Acad. Sci.* 110, 9686–9691. doi:10.1073/pnas.1300195110
- Adams, J.D., Chiang, E.P., Jensen, J.L., 2003. The Influence of Federal Laboratory R&D on Industrial Research. *Rev. Econ. Stat.* 85, 1003–1020. doi:10.1162/003465303772815899
- Alexopoulos, M., 2011. Read All about It!! What Happens Following a Technology Shock? *Am. Econ. Rev.* 1144–1179.
- American Energy Innovation Council, 2010. A Business Plan for America’s Energy Future. Full Report.
- Anadón, L.D., Bosetti, V., Bunn, M., Catenacci, M., Lee, A., 2012. Expert Judgments about RD&D and the Future of Nuclear Energy. *Environ. Sci. Technol.* 46, 11497–11504. doi:10.1021/es300612c
- Anadón, L.D., Bunn, M., Chan, G., Chan, M., Jones, C., Kempener, R., Lee, A., Logar, N., Narayanamurti, V., 2011. Transforming U.S. Energy Innovation. Belfer Center for Science and International Affairs, Harvard Kennedy School.
- Anadón, L.D., Bunn, M., Narayanamurti, V. (Eds.), 2014a. Transforming U.S. Energy Innovation. Cambridge University Press, New York, NY.
- Anadón, L.D., Chan, G., Lee, A., 2014b. Expanding, and improving targeting of, U.S. investment in energy innovation: An analytical approach, in: Anadón, L.D., Bunn, M., Narayanamurti, V. (Eds.), Transforming U.S. Energy Innovation. Cambridge University Press, Cambridge, U.K., and New York, NY, USA, pp. 36–80.

- Anadón, L.D., Nemet, G., Verdolini, E., 2013. The future costs of nuclear power using multiple expert elicitations: effects of RD&D and elicitation design. *Environ. Res. Lett.* 8, 034020. doi:10.1088/1748-9326/8/3/034020
- Angrist, J.D., Pischke, J.-S., 2009. *Mostly harmless econometrics an empiricist's companion*. Princeton University Press, Princeton.
- APS, 2008. *Energy Future: Think Efficiency. How America Can Look Within to Achieve Energy Security and Reduce Global Warming*.
- Arora, A., 1995. Licensing Tacit Knowledge: Intellectual Property Rights And The Market For Know-How. *Econ. Innov. New Technol.* 4, 41–60. doi:10.1080/10438599500000013
- Arora, A., Fosfuri, A., Gambardella, A., 2004. *Markets for technology: the economics of innovation and corporate strategy*. MIT, Cambridge, Mass.; London.
- Arrow, K., 1962. Economic Welfare and the Allocation of Resources for Invention, in: *The Rate and Direction of Inventive Activity: Economic and Social Factors*. pp. 609–626.
- Arthur, W.B., 2009. *The Nature of Technology: What It Is and How It Evolves*. Free Press.
- Baker, E., Bosetti, V., Anadón, L.D., Henrion, M., Reis, L.A., 2014. Future Costs of Key Low-Carbon Energy Technologies: Harmonization and Aggregation of Energy Technology Expert Elicitation Data.
- Baker, E., Chon, H., Keisler, J., 2010. Battery technology for electric and hybrid vehicles: Expert views about prospects for advancement. *Technol. Forecast. Soc. Change* 77, 1139–1146. doi:10.1016/j.techfore.2010.02.005
- Baker, E., Chon, H., Keisler, J., 2009a. Advanced solar R&D: Combining economic analysis with expert elicitations to inform climate policy. *Energy Econ.* 31, S37–S49. doi:10.1016/j.eneco.2007.10.008
- Baker, E., Chon, H., Keisler, J., 2009b. Carbon capture and storage: combining economic analysis with expert elicitations to inform climate policy. *Clim. Change* 96, 379–408. doi:10.1007/s10584-009-9634-y
- Baker, E., Keisler, J.M., 2011. Cellulosic biofuels: Expert views on prospects for advancement. *Energy* 36, 595–605. doi:10.1016/j.energy.2010.09.058
- Black, D.A., Smith, J.A., 2004. How robust is the evidence on the effects of college quality? Evidence from matching. *J. Econom.* 121, 99–124. doi:10.1016/j.jeconom.2003.10.006
- Blanford, G.J., 2009. R&D investment strategy for climate change. *Energy Econ.* 31, S27–S36. doi:10.1016/j.eneco.2008.03.010

- Blei, D., 2010. Introduction to Probabilistic Topic Models.
- Blei, D., Lafferty, J.D., 2007. A correlated topic model of science. *Ann. Appl. Stat.* 1, 17–35.
- Blei, D., Ng, A., Jordan, M., 2003. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3, 993–1022.
- Blind, K., Cremers, K., Mueller, E., 2009. The influence of strategic patenting on companies' patent portfolios. *Res. Policy* 38, 428–436. doi:10.1016/j.respol.2008.12.003
- Bloom, N., Van Reenen, J., 2002. Patents, real options and firm performance. *Econ. J.* 112, C97–C116. doi:10.1111/1468-0297.00022
- Bolinger, M., Wiser, R., 2012. Understanding wind turbine price trends in the U.S. over the past decade. *Energy Policy* 42, 628–641. doi:10.1016/j.enpol.2011.12.036
- Bosetti, V., Catenacci, M., Fiorese, G., Verdolini, E., 2012. The future prospect of PV and CSP solar technologies: An expert elicitation survey. *Energy Policy* 49, 308–317. doi:10.1016/j.enpol.2012.06.024
- Bozeman, B., 2000. Technology transfer and public policy: a review of research and theory. *Res. Policy* 29, 627–655. doi:10.1016/S0048-7333(99)00093-1
- Cannady, C., 2013. Technology licensing and development agreements. Oxford University Press, Oxford [UK]; New York.
- Carrillo, C., Obando Montaña, A.F., Cidrás, J., Díaz-Dorado, E., 2013. Review of power curve modelling for wind turbines. *Renew. Sustain. Energy Rev.* 21, 572–581. doi:10.1016/j.rser.2013.01.012
- Catenacci, M., Verdolini, E., Bosetti, V., Fiorese, G., 2013. Going electric: Expert survey on the future of battery technologies for electric vehicles. *Energy Policy* 61, 403–413. doi:10.1016/j.enpol.2013.06.078
- Chan, G., Anadón, L.D., Chan, M., Lee, A., 2011. Expert elicitation of cost, performance, and RD&D budgets for coal power with CCS. *Energy Procedia* 4, 2685–2692. doi:10.1016/j.egypro.2011.02.169
- Cherp, A., Adenikinju, A., Goldthau, A., Hernandez, F., Hughes, L., Jansen, J., Jewell, J., Olshanskaya, M., Soares de Oliveira, R., Sovacool, B., Vakulenko, S., 2012. Energy and Security, in: *Global Energy Assessment - Toward a Sustainable Future*. Cambridge University Press, Cambridge, UK and New York, NY, USA and the International Institute for Applied Systems Analysis, Laxenburg, Austria, pp. 325–384.
- Cho, A., 2013. Hubs Aim to Reinvent DOE Research Culture. *Science* 340, 914–918. doi:10.1126/science.340.6135.914

- Cohen, L., Noll, R., 1996. The future of the national laboratories. *Proc. Natl. Acad. Sci.* 93, 12678–12685.
- Cooke, R.M., 1991. *Experts in uncertainty: opinion and subjective probability in science.* Oxford University Press, New York.
- Cox, D.R., 1984. *Analysis of survival data, Monographs on statistics and applied probability.* Chapman and Hall, London ; New York.
- CREIA, 2014. *China's Wind Power Installed Capacity Statistics 2013 (2013 中国可再生能源学会风能专业委员会).* China Renewable Energy Industry Association (CREIA), Beijing, China.
- CREIA, CWEA, GWEC, 2014. *Annual Review and Outlook on China Wind Power 2013 (2013 中国风电发展报告).* China Renewable Energy Industry Association (CREIA), Chinese Wind Energy Association (CWEA), Global Wind Energy Council (GWEC), Beijing, China.
- Curtright, A.E., Morgan, M.G., Keith, D.W., 2008a. Expert Assessments of Future Photovoltaic Technologies. *Environ. Sci. Technol.* 42, 9031–9038. doi:10.1021/es8014088
- Curtright, A.E., Morgan, M.G., Keith, D.W., 2008b. Expert Assessments of Future Photovoltaic Technologies. *Environ. Sci. Technol.* 42, 9031–9038. doi:10.1021/es8014088
- Dahan, E., Mendelson, H., 2001. An Extreme-Value Model of Concept Testing. *Manag. Sci.* 47, 102–116. doi:10.1287/mnsc.47.1.102.10666
- Dai, Y., Xue, L., 2015. China's policy initiatives for the development of wind energy technology. *Clim. Policy* 15, 30–57. doi:10.1080/14693062.2014.863549
- Dalkey, N., 1969. An experimental study of group opinion. *Futures* 1, 408–426. doi:10.1016/S0016-3287(69)80025-X
- David, P.A., Hall, B.H., Toole, A.A., 2000. Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Res. Policy* 29, 497–529. doi:10.1016/S0048-7333(99)00087-6
- Davis, G.A., Owens, B., 2003. Optimizing the level of renewable electric R&D expenditures using real options analysis. *Energy Policy* 31, 1589–1608. doi:10.1016/S0301-4215(02)00225-2
- Davis, S.J., Socolow, R.H., 2014. Commitment accounting of CO<sub>2</sub> emissions. *Environ. Res. Lett.* 9, 084018. doi:10.1088/1748-9326/9/8/084018
- Deutch, J., Lester, R., 2004. *Making Technology Work: Applications in Energy and the Environment.* Cambridge University Press, Cambridge.

- DOE, 2012. Quadrennial Technology Review. U.S. Department of Energy.
- DOE External Expert Peer Review Panel, 2006. Report of the External Expert Peer Review Panel: DOE Benefits Forecasts.
- DOE Office of Energy Efficiency and Renewable Energy, 2014. Energy Innovation Portal.
- DOE Office of Energy Efficiency and Renewable Energy, 2012. FY 2009 - 2012 DOE Technology Transfer Data (Internal Database).
- DOE Office of Science, 2013. The Office of Science Laboratories.
- DOE Office of Scientific and Technical Information, 2014a. Department of Energy National Laboratories.
- DOE Office of Scientific and Technical Information, 2014b. DOepatents.
- DOE Technology Transfer Working Group, 2013. Licensing Guide and Sample License.
- Drivas, K., Lei, Z., Wright, B., 2014. Academic Patent Licenses: Roadblocks or Signposts for Nonlicensee Cumulative Innovation? SSRN Work. Pap. 2489231.
- EIA, 2009. EIA-Annual Energy Outlook 2009 (No. DOE/EIA-0383(2009)). U.S. Energy Information Administration.
- Executive Office of the President, 2013. The President's Climate Action Plan.
- Federal Laboratory Consortium for Technology Transfer, 2011. Technology Transfer Desk Reference: A Comprehensive Guide to Technology Transfer.
- Feinerer, I., Hornik, K., Meyer, D., 2008. Text mining infrastructure in R. *J. Stat. Softw.* 25, 1–54.
- Figueira, J., Greco, S., Ehrgott, M., 2005. Multiple criteria decision analysis state of the art surveys. Springer, New York.
- Fiorese, G., Catenacci, M., Verdolini, E., Bosetti, V., 2013. Advanced biofuels: Future perspectives from an expert elicitation survey. *Energy Policy* 56, 293–311. doi:10.1016/j.enpol.2012.12.061
- Fishbone, L.G., Abilock, H., 1981. Markal, a linear-programming model for energy systems analysis: Technical description of the bnl version. *Int. J. Energy Res.* 5, 353–375. doi:10.1002/er.4440050406
- Fleming, L., 2001. Recombinant Uncertainty in Technological Search. *Manag. Sci.* 47, 117–132. doi:10.1287/mnsc.47.1.117.10671
- Furman, J.L., Stern, S., 2011. Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research. *Am. Econ. Rev.* 101, 1933–1963. doi:10.1257/aer.101.5.1933

- Galasso, A., Schankerman, M., 2013. Patents and Cumulative Innovation: Causal Evidence from the Courts. Rotman Sch. Manag. Work. Pap.
- Gallagher, K.S., Anadón, L.D., 2014. DOE Budget Authority for Energy Research, Development, & Demonstration Database. Energy Technology Innovation Policy research group, Belfer Center for Science and International Affairs, Harvard Kennedy School.
- Gallagher, K.S., Anadón, L.D., 2013. DOE Budget Authority for Energy Research, Development, & Demonstration Database. Energy Technology Innovation Policy research group, Belfer Center for Science and International Affairs, Harvard Kennedy School.
- Gans, J.S., Stern, S., 2010. Is there a market for ideas? *Ind. Corp. Change* 19, 805–837. doi:10.1093/icc/dtq023
- Geum, Y., Park, Y., 2013. The state-of-the-art of public-sector technology roadmaps: A taxonomical approach to energy technology roadmaps. *Sci. Public Policy* 40, 327–339. doi:10.1093/scipol/scs108
- Google, 2014. Google Patents.
- Goolsbee, A., 1998. Does Government R&D Policy Mainly Benefit Scientists and Engineers? *Am. Econ. Rev.* 88, 298–302.
- Gosens, J., Lu, Y., 2013. From lagging to leading? Technological innovation systems in emerging economies and the case of Chinese wind power. *Energy Policy* 60, 234–250. doi:10.1016/j.enpol.2013.05.027
- Granot, D., Zuckerman, D., 1991. Optimal Sequencing and Resource Allocation in Research and Development Projects. *Manag. Sci.* 37, 140–156. doi:10.1287/mnsc.37.2.140
- Greening, L.A., Bernow, S., 2004. Design of coordinated energy and environmental policies: use of multi-criteria decision-making. *Energy Policy* 32, 721–735. doi:10.1016/j.enpol.2003.08.017
- Green, J., Scotchmer, S., 1995. On the division of profit in sequential innovation. *RAND J. Econ.* 26, 20–33.
- Griliches, Z., 1992. The Search for R&D Spillovers. *Scand. J. Econ.* 94, S29–47.
- Grimmer, J., Stewart, B.M., 2013. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Polit. Anal.* 21, 267–297. doi:10.1093/pan/mps028
- Gruen, B., Hornik, K., 2011. topicmodels: Topic models. R package version 0.0-8.

- Hall, B., 1996. The private and social returns to research and development, in: Smith, B., Barfield, C. (Eds.), *Technology, R&D, and the Economy*. The Brookings Institution.
- Hall, B.H., Mairesse, J., Mohnen, P., 2010. Measuring the Returns to R&D, in: *Handbook of the Economics of Innovation*. Elsevier, pp. 1033–1082.
- Harhoff, D., Scherer, F.M., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. *Res. Policy* 32, 1343–1363. doi:10.1016/S0048-7333(02)00124-5
- Hausman, N., 2010. Effects of university innovation on local economic growth and entrepreneurship. Harvard University, Department of Economics, Job Market Paper.
- Heckman, J., Navarro-Lozano, S., 2004. Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. *Rev. Econ. Stat.* 86, 30–57. doi:10.1162/003465304323023660
- He, G., Morse, R., 2013. Addressing Carbon Offsetters' Paradox: Lessons from Chinese Wind CDM. *Energy Policy* 63, 1051–1055.
- He, G., Morse, R., 2010. Making Carbon Offsets Work in the Developing World: Lessons from the Chinese Wind Controversy (No. Program on Energy and Sustainable Development Working Paper #90). Stanford University Freeman Spogli Institute for International Studies.
- Hellmann, T., 2007. The role of patents for bridging the science to market gap. *J. Econ. Behav. Organ.* 63, 624–647. doi:10.1016/j.jebo.2006.05.013
- Henderson, R., Jaffe, A.B., Trajtenberg, M., 1998. Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988. *Rev. Econ. Stat.* 80, 119–127. doi:10.1162/003465398557221
- Hirabayashi, J., 2003. Revisiting the USPTO Concordance Between the US Patent Classification and the Standard Industrial Classification Systems. Presented at the WIPO-OECD Workshop on Statistics in the Patent Field, Geneva, Switzerland.
- Hoberg, G., Phillips, G., 2010. Dynamic Text-Based Industries and Endogenous Product Differentiation. NBER Work. Pap. Ser.
- Holland, P.W., 1986. Statistics and Causal Inference. *J. Am. Stat. Assoc.* 81, 945–960. doi:10.1080/01621459.1986.10478354
- Hopkins, D., King, G., 2010. A Method of Automated Nonparametric Content Analysis for Social Science. *Am. J. Polit. Sci.* 229–247.
- Hornik, K., 2007. Snowball: Snowball Stemmers. R package version 0.0-1.

- Hsu, Y.-G., Tzeng, G.-H., Shyu, J.Z., 2003. Fuzzy multiple criteria selection of government-sponsored frontier technology R&D projects. *R Manag.* 33, 539–551. doi:10.1111/1467-9310.00315
- Huaxia, 2013. Chinese Wind Power Market and Wind Farm Construction Report (2013 年中国风电市场暨已建在建风电场项目研究报告). Huaxia Wind Ltd., Beijing, China.
- Huaxia, 2011. Wind Turbine Manufacturer Statistics 2011 (2011 年最新风电整机厂商统计). Huaxia Wind Ltd., Beijing, China.
- Iacus, S.M., King, G., Porro, G., 2012. Causal Inference without Balance Checking: Coarsened Exact Matching. *Polit. Anal.* 20, 1–24. doi:10.1093/pan/mpr013
- IEA, 2014. CO<sub>2</sub> Emissions from Fuel Combustion (No. 2014 Edition). Paris, France.
- IEA, 2013. RD&D Database. Organization for Economic Cooperation and Development, Paris, France.
- IEA, 2011. World Energy Outlook 2011. IEA, International Energy Agency: OECD, Paris.
- IGES, 2015a. CDM Project Database. Institute for Global Environmental Strategies (IGES), Hayama, Japan.
- IGES, 2015b. CDM Investment Analysis Database. Institute for Global Environmental Strategies (IGES), Hayama, Japan.
- Iman, R.L., Conover, W.J., 1982. A Distribution-Free Approach to Inducing Rank Order Correlation Among Input Variables. *Commun. Stat. - Simul. Comput.* 11, 311–334.
- Imbens, G.W., 2004. Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *Rev. Econ. Stat.* 86, 4–29. doi:10.1162/003465304323023651
- INL, 2006. Wind Turbine Power Curve Database. Idaho National Labs (INL), Idaho Falls, Idaho, US.
- IPCC, 2014. Climate change 2014: Mitigation of climate change: Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York.
- Jaffe, A., 2012. Technology Policy and Climate Change, The Next Round of Climate Economics & Policy Research. Washington D.C.
- Jaffe, A.B., Newell, R.G., Stavins, R.N., 2005. A tale of two market failures: Technology and environmental policy. *Ecol. Econ.* 54, 164–174. doi:10.1016/j.ecolecon.2004.12.027



- Jaffe, A.B., Trajtenberg, M., 1999. International Knowledge Flows: Evidence From Patent Citations. *Econ. Innov. New Technol.* 8, 105–136. doi:10.1080/10438599900000006
- Jaffe, A.B., Trajtenberg, M., 1996. Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proc. Natl. Acad. Sci.* 93, 12671–12677.
- Jaffe, A.B., Trajtenberg, M., Fogarty, M.S., 2000. Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *Am. Econ. Rev.* 90, 215–218. doi:10.1257/aer.90.2.215
- Jaffe, A., Lerner, J., 2001. Reinventing Public R&D: Patent Policy and the Commercialization of National Laboratory Technologies. *RAND J. Econ.* 167–198.
- Jaffe, A., Trajtenberg, M. (Eds.), 2002. *Patents, Citations, and Innovations*. MIT Press.
- Jaffe, A., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Q. J. Econ.* 557–598.
- Jenni, K.E., Baker, E.D., Nemet, G.F., 2013. Expert elicitations of energy penalties for carbon capture technologies. *Int. J. Greenh. Gas Control* 12, 136–145. doi:10.1016/j.ijggc.2012.11.022
- Kaplan, S., Vakili, K., 2014. The Double-Edged Sword of Recombination in Breakthrough Innovation. *Strateg. Manag. J.* doi:10.1002/smj.2294
- Keefer, D.L., 1991. Resource Allocation Models with Risk Aversion and Probabilistic Dependence: Offshore Oil and Gas Bidding. *Manag. Sci.* 37, 377–395. doi:10.1287/mnsc.37.4.377
- Keith, D.W., 1996. When is it appropriate to combine expert judgments?: An editorial essay. *Clim. Change* 33, 139–143. doi:10.1007/BF00140244
- King, G., Pan, J., Roberts, M.E., 2013. How Censorship in China Allows Government Criticism but Silences Collective Expression. *Am. Polit. Sci. Rev.* 107, 326–343. doi:10.1017/S0003055413000014
- LBNL, Innovation and Partnerships Office, 2014. *Licensing Overview*.
- Lema, A., Lema, R., 2013. Technology transfer in the clean development mechanism: Insights from wind power. *Glob. Environ. Change* 23, 301–313. doi:10.1016/j.gloenvcha.2012.10.010
- Lewis, J.I., 2013. *Green Innovation in China: China's wind power industry and the global transition to a low-carbon economy*, Contemporary Asia in the world. Columbia University Press, New York.

- Lewis, J.I., 2010a. The evolving role of carbon finance in promoting renewable energy development in China. *Energy Policy* 38, 2875–2886. doi:10.1016/j.enpol.2010.01.020
- Lewis, J.I., 2010b. The evolving role of carbon finance in promoting renewable energy development in China. *Energy Policy* 38, 2875–2886. doi:10.1016/j.enpol.2010.01.020
- Liberatore, M.J., Titus, G.J., 1983. The Practice of Management Science in R&D Project Management. *Manag. Sci.* 29, 962–974. doi:10.1287/mnsc.29.8.962
- Loch, C.H., Terwiesch, C., Thomke, S., 2001. Parallel and Sequential Testing of Design Alternatives. *Manag. Sci.* 47, 663–678. doi:10.1287/mnsc.47.5.663.10480
- Logar, N., Narayanamurti, V., Anadón, L.D., 2014. Reforming U.S. Energy Innovation Institutions: Maximizing the Return on Investment, in: Anadón, L.D., Bunn, M., Narayanamurti, V. (Eds.), *Transforming U.S. Energy Innovation*. pp. 81–124.
- Loulou, R., Goldsten, G., Noble, K., n.d. Documentation for the MARKAL Family of Models.
- Margolis, R.M., Kammen, D.M., 1999. Evidence of under-investment in energy R&D in the United States and the impact of Federal policy. *Energy Policy* 27, 575–584. doi:10.1016/S0301-4215(99)00053-1
- Martino, J.P., 1987. Using precursors as leading indicators of technological change. *Technol. Forecast. Soc. Change* 32, 341–360. doi:10.1016/0040-1625(87)90051-5
- McElroy, M.B., Lu, X., Nielsen, C.P., Wang, Y., 2009. Potential for Wind-Generated Electricity in China. *Science* 325, 1378–1380. doi:10.1126/science.1175706
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21, 239–245. doi:10.1080/00401706.1979.10489755
- McNerney, J., Doyne Farmer, J., Trancik, J.E., 2011. Historical costs of coal-fired electricity and implications for the future. *Energy Policy* 39, 3042–3054. doi:10.1016/j.enpol.2011.01.037
- Merton, R.K., 1973. *The sociology of science: theoretical and empirical investigations*. University of Chicago Press, Chicago.
- Michaelowa, A., 2012. Strengths and weaknesses of the CDM in comparison with new and emerging market mechanisms (No. Paper No. 2), *CDM Policy Dialogue*.
- Michaelowa, A., Stronzik, M., Eckermann, F., Hunt, A., 2003. Transaction costs of the Kyoto Mechanisms. *Clim. Policy* 3, 261–278. doi:10.3763/cpol.2003.0332

- Moniz, E., 2013. Testimony of Secretary Ernest Moniz U.S. Department of Energy Before the House Committee on Energy and Commerce Subcommittee on Energy and Power.
- Morgan, M.G., 2014. Use (and abuse) of expert elicitation in support of decision making for public policy. *Proc. Natl. Acad. Sci.* 111, 7176–7184. doi:10.1073/pnas.1319946111
- Morgan, M.G., Henrion, M., 1998. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press, Cambridge.
- Morgan, M.G., Keith, D.W., 1995. Subjective Judgments by Climate Experts. *Environ. Sci. Technol.* 29, 468A–476A. doi:10.1021/es00010a753
- Mostellar, F., Wallace, D., 1964. *Inference and Disputed Authorship*. University of Chicago Press.
- Mowery, D.C., Sampat, B.N., Ziedonis, A.A., 2002. Learning to Patent: Institutional Experience, Learning, and the Characteristics of U.S. University Patents After the Bayh-Dole Act, 1981-1992. *Manag. Sci.* 48, 73–89. doi:10.1287/mnsc.48.1.73.14278
- Mowery, D., Richard, N., Bhaven, S., Ziedonis, A., 2001. The growth of patenting and licensing by U.S. universities: an assessment of the effects of the Bayh-Dole act of 1980. *Res. Policy* 99–119.
- Murray, F., Stern, S., 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? *J. Econ. Behav. Organ.* 63, 648–687. doi:10.1016/j.jebo.2006.05.017
- NASA, 2005. NASA SSE monthly average wind data at one-degree resolution of the world. U.S. National Aeronautics and Space Administration (NASA), Surface meteorology and Solar Energy (SSE).
- National Institute of Standards and Technology, Anderson, G., 2011. *Federal Technology Transfer Data*.
- National Science Board, 2014. *Science and Engineering Indicators 2014 (No. NSB 14-01)*. National Science Foundation, Arlington, VA.
- NBS, 2013. *China Energy Statistical Yearbook*. Compiled by the Energy Statistics Division of the National Bureau of Statistics (NBS), Published by China Statistics Press, Beijing, China.
- NCEP, 2004. *Ending the Energy Stalemate. A Bipartisan Strategy to Meet America's Energy Challenges*. The National Commission on Energy Policy.
- NEA, 2015. *2014 Wind Power Industry Monitoring*. National Energy Agency (NEA), Beijing, China.

- Nelson, R., 1959. The Simple Economics of Basic Scientific Research. *J. Polit. Econ.* 67, 297–306.
- Nemet, G.F., 2006. Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy* 34, 3218–3232. doi:10.1016/j.enpol.2005.06.020
- Nemet, G.F., Kammen, D.M., 2007. U.S. energy research and development: Declining investment, increasing need, and the feasibility of expansion. *Energy Policy* 35, 746–755. doi:10.1016/j.enpol.2005.12.012
- NRC, 2007a. Prospective Evaluation of Applied Energy Research and Development at DOE (Phase Two). National Research Council, Washington, D.C.
- NRC, 2007b. Prospective evaluation of applied energy research and development at DOE (Phase two), National Research Council (U.S.). National Academies Press, Washington, D.C.
- NRC, 2001. Energy research at DOE, was it worth it?: energy efficiency and fossil energy research 1978 to 2000, National Research Council (U.S.). National Academy Press, Washington, D.C.
- Oppenheimer, M., O'Neill, B.C., Webster, M., Agrawala, S., 2007. The Limits of Consensus. *Science* 317, 1505–1506. doi:10.1126/science.1144831
- Pakes, A., 1986. Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica* 54, 755–784.
- Parson, E., 2002. The Technology Assessment Approach to Climate Change. *Issues Sci. Technol.* 18, 65–72.
- Parson, E., 1998. The Montreal Protocol: The First Adaptive Global Environmental Regime?, in: *Protecting the Ozone Layer: Lessons, Models, and Prospects*. Kluwer, Boston, pp. 127–134.
- Partha, D., David, P.A., 1994. Toward a new economics of science. *Res. Policy* 23, 487–521. doi:10.1016/0048-7333(94)01002-1
- PCAST, 2010. Accelerating the Pace of Change in Energy Technologies through an Integrated Federal Energy Policy: Report to the President. DIANE Publishing.
- PCAST, 1997. Federal Energy Research and Development for the Challenges of the Twenty-First Century. Executive Office of the President, Washington D.C.
- PNNL, Technology Transfer, 2014. Our Licensing Guidelines. Pacific Northwest National Lab.
- Pugh, G., Clarke, L., Marlay, R., Kyle, P., Wise, M., McJeon, H., Chan, G., 2011. Energy R&D portfolio analysis based on climate change mitigation. *Energy Econ.* 33, 634–643. doi:10.1016/j.eneco.2010.11.007

- Quinn, K.M., Monroe, B.L., Colaresi, M., Crespin, M.H., Radev, D.R., 2010. How to analyze political attention with minimal assumptions and costs. *Am. J. Polit. Sci.* 54, 209–228.
- Raiffa, H., 1968. *Decision Analysis*. Addison-Wesley, Reading, MA.
- Ranson, M., Stavins, R.N., 2015. Linkage of greenhouse gas emissions trading systems: learning from experience. *Clim. Policy* 1–17. doi:10.1080/14693062.2014.997658
- Rooksby, J., 2013. Innovation and Litigation: Tensions Between Universities and Patents and How to Fix Them. *Yale J. Law Technol.* 15, 312–404.
- Roper, A.T., Cunningham, S.W., Porter, A.L., Mason, T.W., Rossini, F.A., Banks, 2011. *Forecasting and management of technology*, 2nd ed. ed. John Wiley & Sons, Hoboken, N.J.
- Rosenbaum, P., 2005. Observational Study, in: *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Hoboken, N.J.
- Rosenbaum, P., 2002. *Observational Studies*, 2nd ed. Springer, New York, NY.
- Rosenberg, N., 1982. *Inside the black box: technology and economics*. Cambridge University Press, Cambridge ; New York.
- Rubin, D., 2006. *Matched Sampling for Causal Effects*. Cambridge University Press, Cambridge, UK.
- Ruegg, R., Feller, I., 2003. *A Toolkit for Evaluating Public R&D Investment: Models, Methods, and Findings from ATP's First Decade (No. NIST GCR 03-857)*. National Institute of Standards and Technology.
- Santen, N., Anadón, L.D., 2014. *Electricity Technology Investments Under Solar RD&D Uncertainty: How Interim Learning and Adaptation Affect the Optimal Decision Strategy (No. Discussion Paper #2014-10)*. Belfer Center for Science and International Affairs, Harvard Kennedy School.
- Santiago, L.P., Vakili, P., 2005. On the Value of Flexibility in R&D Projects. *Manag. Sci.* 51, 1206–1218. doi:10.1287/mnsc.1050.0387
- Scherer, F.M., 1982. Inter-Industry Technology Flows and Productivity Growth. *Rev. Econ. Stat.* 627–634.
- Scherer, F.M., Harhoff, D., 2000. Technology policy for a world of skew-distributed outcomes. *Res. Policy* 29, 559–566. doi:10.1016/S0048-7333(99)00089-X
- Schock, R.N., Fulkerson, W., Brown, M.L., San Martin, R.L., Greene, D.L., Edmonds, J., 1999. How much is energy research & development worth as insurance? *Annu. Rev. Energy Environ.* 24, 487–512. doi:10.1146/annurev.energy.24.1.487

- Scotchmer, S., 1991. Standing on the Shoulders of Giants: Cumulative Research and the Patent Law. *J. Econ. Perspect.* 5, 29–41. doi:10.1257/jep.5.1.29
- Shi, P., 2003. Chinese Wind Farm Capacity Statistics 2002 (2002 年中国风电场装机容量统计). Beijing, China.
- Silverman, B.G., 1981. Project Appraisal Methodology: A Multidimensional R&D Benefit/Cost Assessment Tool. *Manag. Sci.* 27, 802–821. doi:10.1287/mnsc.27.7.802
- Singh, J., Agrawal, A., 2011. Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires. *Manag. Sci.* 57, 129–150. doi:10.1287/mnsc.1100.1253
- SPC, 2002. Interim Rules of Electrical Engineering Retrofit Projects. State Power Corporation of China.
- Stange, K., 2011. A Longitudinal Analysis of the Relationship Between Fertility Timing and Schooling. *Demography* 48, 931–956. doi:10.1007/s13524-011-0050-3
- Stepp, M., Pool, S., Loris, N., Spencer, J., 2013. Turning the Page: Reimagining the National Labs in the 21st Century Innovation Economy. The Information Technology and Innovation Foundation, The Center for American Progress, and The Heritage Foundation.
- Stevenson-Wydler Technology Innovation Act of 1980, 1980.
- Stewart, T.J., 1991. A Multi-criteria Decision Support System for R&D Project Selection. *J. Oper. Res. Soc.* 42, 17–26. doi:10.1057/jors.1991.3
- Stua, M., 2013. Evidence of the clean development mechanism impact on the Chinese electric power system's low-carbon transition. *Energy Policy* 62, 1309–1319. doi:10.1016/j.enpol.2013.07.039
- Tang, T., Popp, D., 2014. The Learning Process and Technological Change in Wind Power: Evidence from China's CDM Wind Projects. NBER Work. Pap. Ser. 19921.
- Terwiesch, C., Loch, C.H., 1999. Measuring the Effectiveness of Overlapping Development Activities. *Manag. Sci.* 45, 455–465. doi:10.1287/mnsc.45.4.455
- The Economist, 2010. Carbon Markets: The smoking greenhouse gun 22–23.
- The White House, 2014. The President's Management Agenda for Fiscal Year 2015: Opportunity for All: Creating a 21st Century Government.
- The White House, 2011. Presidential Memorandum - Accelerating Technology Transfer and Commercialization of Federal Research in Support of High-Growth Businesses. Office of the Press Secretary.
- Thompson, P., Fox-Kean, M., 2005. Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. *Am. Econ. Rev.* 95, 450–460.

- Thomson Reuters, 2012. CER Prices and Volumes. Thomson Reuters Eikon - Research and Forecasts for Energy, New York, NY.
- Turner, A., 2004. PNNL's Body Scanner Garners Federal Commercialization Award. Pacific Northwest National Lab.
- UNEP Risoe Center, 2013. UNEP Risoe CDM/JI Pipeline Analysis and Database.
- UNFCCC, 2014. Draft Large-scale Consolidated Methodology ACM0002: Grid-connected electricity generation from renewable sources. United Nations Framework Convention on Climate Change (UNFCCC).
- UNFCCC, 2012a. Views on the new market-based mechanism - Submissions from Parties. 17th COP at Bonn. United Nations Framework Convention on Climate Change (UNFCCC) Ad Hoc Working Group on Long-term Cooperative Action under the Convention, Bonn, Germany.
- UNFCCC, 2012b. CDM Methodological Tool 01: Tool for the demonstration and assessment of additionality. United Nations Framework Convention on Climate Change (UNFCCC), Bonn, Germany.
- U.S. Government Accountability Office, 2009. Technology Transfer: Clearer Priorities and Greater Use of Innovative Approaches Could Increase the Effectiveness of Technology Transfer at Department of Energy Laboratories (No. GAO-09-548).
- U.S. Patent and Trademark Office, 2014. Patent Full-Text Databases.
- Venugopalan, S., Rai, V., 2014. Topic based classification and pattern identification in patents. Technol. Forecast. Soc. Change forthcoming. doi:10.1016/j.techfore.2014.10.006
- Weitzman, M., 1998. Recombinant Growth. Q. J. Econ. 113, 331–360.
- Westwick, P., 2003. The National Labs: Science in an American System, 1947-1974. Harvard University Press.
- WHO, UNDP, 2009. The Energy Access Situation in Developing Countries. World Health Organization and the United Nations Development Programme, New York.
- Wiesenthal, T., Dowling, P., Morbee, J., Thiel, C., Schade, B., Russ, P., Simoes, S., Peteves, S., Schoots, K., Londo, M., 2012. Technology Learning Curves for Energy Policy Support, JRC Scientific and Policy Reports.
- Williams, H., 2013. Intellectual Property Rights and Innovation: Evidence from the Human Genome. J. Polit. Econ. 121, 1–27.
- Wright, B.D., Drivas, K., Lei, Z., Merrill, S.A., 2014. Technology transfer: Industry-funded academic inventions boost innovation. Nature 507, 297–299. doi:10.1038/507297a

Zhang, J., Wang, C., 2011. Co-benefits and additionality of the clean development mechanism: An empirical analysis. *J. Environ. Econ. Manag.* 62, 140–154. doi:10.1016/j.jeem.2011.03.003



# Appendix A

## Appendix to Chapter 1

### A.1 Expert Elicitation Protocol

This appendix provides additional details on the protocol we developed and followed to conduct a suite of energy technology expert elicitations in 2009-2010. This protocol is described in greater detail in Anadón, et al. (2014b), which forms the basis for this appendix.

#### A.1.1 Technology Selection

We designed and conducted expert elicitations with more than 100 experts<sup>44</sup> on seven technology areas: fuels and electricity from biomass (bioenergy, for short); different types of utility scale energy storage (or storage); residential, commercial, and utility scale photovoltaic technologies (solar); efficiency in commercial buildings (buildings); nuclear power from Generation III/III+, Generation IV, and small and medium reactors (nuclear); coal and natural gas electricity production with and without carbon capture and storage (fossil); and vehicle technologies (vehicles). We were unable to include other

---

<sup>44</sup> 108 experts gave total and partial answers to the different surveys: 30 participated in the nuclear survey, 25 in the storage survey, 12 in the bioenergy survey, 9 in the buildings survey, 9 in the vehicles survey, 12 in the fossil energy survey, and 11 in the PV survey.

energy technologies in DOE’s investment portfolio – notably wind power, geothermal power, concentrated solar power, advanced lighting, and industrial energy efficiency – in this study owing to limited resources. Due to the design of the buildings elicitation, we are not able to quantify uncertainty in building technology parameters in MARKAL, and therefore the results of the buildings elicitation are not included in our analysis in this chapter.

### **A.1.2 Expert Selection**

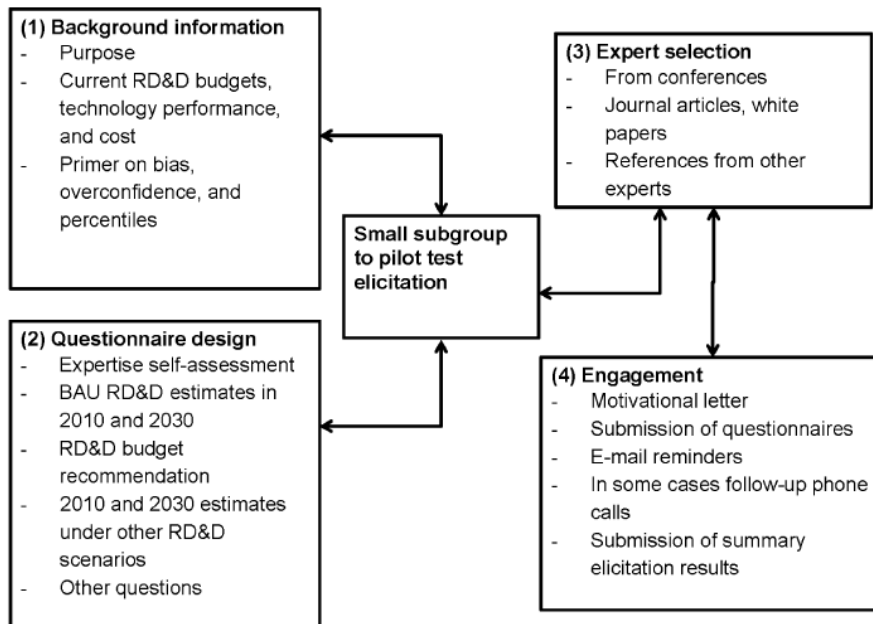
We invited leading researchers and practitioners in each of the seven fields from academia, the private sector – including small and large firms – and the U.S. national laboratories to give us their input. We asked these experts to estimate technical cost and performance metrics in 2030 under a BAU scenario for federal R&D funding for that technology, to recommend the level of annual R&D funding and allocation that would be necessary to increase the commercial viability of the technologies in question, and to revise their 2030 technology projections under different hypothetical budgets.

### **A.1.3 Expert Elicitation Protocol**

Our seven elicitations were carried out using a consistent methodology. We examined four energy supply technology areas (fossil, nuclear, solar, biofuels), two energy demand technology areas (vehicles, buildings), and one enabling energy technology area (energy storage). Figure A.1 illustrates the different tasks that were involved in the design and execution of each of the seven elicitations. Designing and fielding each elicitation took between four and eight months. The first phase involved conducting extensive research into each technology, which we summarized for expert participants in a background information section.

### *Elicitation Design*

### *Expert Selection & Engagement*



Adapted from Anadón et al. (2014b).

**Figure A.1:** *Schematic of the Protocol for Expert Elicitations*

In the second phase, we designed the questionnaires for each technology. The questionnaire first asked experts to self-assess their expertise in a wide range of technology subfields; we used this information to explore potential biases in experts' technical assessments. The questionnaire also asked experts to recommend an aggregate level of R&D funding, and then to estimate cost and performance metrics once in 2010 and then again in 2030 under four scenarios of R&D funding: BAU, half of the recommended budget, the recommended budget, and 10 times the recommended budget. The results of this suite of conditional cost and performance metrics allowed us to explore the sensitivity of the cost and performance estimates to different public R&D investments. A small subgroup of experts (typically two to three) was then used to test and refine the first elicitation draft to increase our confidence that each elicitation

instrument draft would be correctly interpreted by the experts (and that it would provide information in the right form to be of use in the next stages of the method). This second phase of fielding the final elicitation instrument took between two and three months.

In the third phase, we collected the names of experts from a range of sources by examining the peer-reviewed literature, national laboratory programs, university research programs, conference participation, and referrals from other experts and our own program's advisory board. The participant pool we assembled covered a range of perspectives and included technical experts from the private sector, academia, and the national laboratories.

In the fourth phase, we engaged experts through email solicitations for participation. For willing participants, survey responses were conducted by mailed hard copy or through an online platform. In many cases it was necessary to send reminders and hold follow up phone calls to clarify specific questions from the participants. On average, experts invested between two and five hours to complete our study's elicitations, not including the interaction with researchers in the cases in which it took place. Designing and fielding an expert elicitation is a very labor-intensive process for both study designers and participating experts.

We now turn to describing in more detail the structure of the elicitation instruments. The elicitations began an extensive background section divided into four subsections: (i) a summary of the purpose of the elicitation (reminding experts of the rationale explained in their invitation via e-mail and phone) and a note encouraging experts to contact researchers at all times to answer questions; (ii) a technology primer of material on current technology cost and performance, fuel costs if applicable, a summary of current government R&D investments in the technology area, and future cost projections found

in the literature; (iii) a short tutorial on bias and overconfidence, which included a graphical example on expert's overconfidence estimating the speed of light, and instructions about how to reduce overconfidence, including the provision of the expert's 10th and 90th percentile estimates before the 50th percentile estimate, and including the importance of imagining alternate scenarios wherein the true value is outside the ranges the expert provided (which should lead to expert to broaden his/her estimates); and (iv) an explanation of percentiles, including text using language like that used in the remainder of the elicitation, and a graphical representation of interpreting percentiles (Baker et al., 2014).

The background material served to ensure that all experts had the most recent literature fresh in their minds and to encourage them to think consistently about the variables that we would ask them to estimate, which included costs, efficiency, and government R&D investments and programs. The elicitation device was then divided into four or five sections of questions.

In Part 1, experts were asked to assess their expertise on specific technologies, components, and ancillary topics such as feedstocks, specific technology areas, materials, products, and enabling technologies.

In Part 2, experts were asked to identify commercially viable technologies in 2010, as well as the projected cost and performance that would result from a continuation of 2010 federal funding and private investments through 2030, assuming no new government policies are implemented. This scenario was defined as the BAU scenario.<sup>45</sup>

---

<sup>45</sup> The BAU scenario in this work includes all current deployment policies modeled in the Energy Information Administration's (EIA) Annual Energy Outlook 2010 (EIA, 2010a). Anadón et al. (2011) summarizes the status quo policies in the BAU scenario and other future policies categorized into power standards, building codes, transportation policies, and climate policies. It also describes the scenarios used to estimate the impact of oil and gas prices.

In Part 3, experts were asked to recommend a total annual federal R&D budget for the technology area in question. The experts were asked to allocate their recommended budget among basic research, applied research, development, and demonstration investments for specific technologies within the general class of technologies being assessed (e.g., oxy-fired carbon capture technology was one specific technology in the fossil elicitation). These questions asked experts to visually allocate R&D funds into different technology areas and development stages. Experts were also asked to indicate potential coordination with other areas of energy technology research, as well as industries that could provide “spillover” innovations. In most surveys, experts were also asked to provide their insights into the technological hurdles that could be overcome by research in the areas where they recommended the largest investments, and to recommend research areas for cooperation with other countries.

In Part 4, experts were asked to update their BAU 2030 technology cost and performance estimates under three scenarios of R&D funding. These scenarios were defined in relation to the level of R&D recommended by the experts. First experts estimated 2030 cost and performance metrics assuming their recommended annual federal R&D budget was implemented and held constant over the next 20 years. Then, experts updated their 2030 estimates under two additional R&D scenarios: a 50% proportional reduction in their recommended budget and a 10-fold proportional increase in their recommended budget. Having experts re-estimate future technology cost and performance was a crucial part of this analysis because it allowed us to estimate the sensitivity of technology outcomes to R&D funding. In the online elicitations, technology cost questions under different R&D scenarios were visualized in one graph enabling and encouraging respondents to adjust their answers as they were completing the elicitation.

In Part 5, which was not formally a part of all the elicitations we conducted, experts were asked to estimate deployment levels that could be achieved under the four R&D budget scenarios. These elicitations also asked experts to think through deployment policies that would contribute to commercializing novel energy technologies. In those surveys where deployment was not examined in a separate section, experts were asked similar questions about deployment in Parts 2 and 4.

Four of the elicitations (the bioenergy, fossil, storage, and vehicles surveys) were conducted using a written device, which was mailed to participants. The remaining three elicitations (the nuclear, buildings, and solar surveys) were conducted online. Online elicitations improve the ability of experts to modify their answers and to visualize them as they input their estimates. We included several graphics that allowed the experts to see the uncertainty ranges they specified as well as their estimates of cost and performance under different budget scenarios alongside each other. Online elicitations also accelerated the data collection and analysis process, which would be beneficial for future elicitations conducted on a more frequent or broader scale.

At the end of the elicitations, all experts were provided with a written summary of the responses of all participating experts, with the ability to change theirs. For the elicitation on nuclear energy, we also convened a workshop of the participants in the elicitation we conducted and the participants from a similar expert elicitation conducted by the Fondazione Eni Enrico Mattei (FEEM). These experts were given the possibility of revising their responses in private after each workshop session. The details of the workshop and the lessons learned are described in Anadón et al., 2012 (Anadón et al., 2012).

For reference, the nuclear energy expert elicitation is publicly available online at [https://erd3.cloudapp.net/nuclear\\_energy](https://erd3.cloudapp.net/nuclear_energy).

#### **A.1.4 Qualitative Reviews of Elicitation Results**

We complemented the seven elicitations with qualitative interviews (two to six per elicitation) in which we presented the elicitation results to a set of 23 additional experts who were not involved in the first round of elicitations but had expertise managing R&D budgets for each technology area and experience thinking about investments on a range of technology projects. In these conversations, which lasted from one to two hours, we showed the program experts the technology area experts' recommended budgets and technology cost and performance parameter estimates. Their affiliations included DOE programs, venture capital firms, and U.S. Congressional committees. These qualitative reviews helped us interpret the elicitation results and served to expand expert input in our work. These additional experts gave their views on how to synthesize the entire set of results from each of the elicitations and helped identify a representative “middle” expert in each survey (see Section 1.2.2).

### **A.2 Correlation of Future Technology Cost Improvements**

In Section 1.1.2, we describe the requirements for and the process we used to construct a correlation matrix relating the 2030 costs of the 25 technologies considered in this chapter. Table A.1 displays this correlation matrix, expressed as a Spearman correlation matrix.



Table A.1: Estimated Cross-Technology Spearman Correlations

index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
tech code	GTC	DTC	JTC	ETC	COL	GAS	CCS	GCC	GBC	DBC	CAS	BLI	BNS	FLO	BEV	CAR	HYB	PEV	FCV	THR	FOR	MOD	FVR	FVC	FVU	
GTC	1.00	0.95	0.95	0.70	0.60	0.60	0.70	0.70																		
DTC	0.95	1.00	0.95	0.70	0.60	0.60	0.70	0.70																		
JTC	0.95	0.95	1.00	0.70	0.60	0.60	0.70	0.70																		
ETC	0.70	0.70	0.70	1.00	0.70	0.70	0.65	0.65																		
COL	0.60	0.60	0.60	0.70	1.00	0.70	0.90	0.70																		
GAS	0.60	0.60	0.60	0.70	0.70	1.00	0.70	0.90																		
CCS	0.70	0.70	0.70	0.65	0.90	0.70	1.00	0.80																		
GCC	0.70	0.70	0.70	0.65	0.70	0.90	0.80	1.00																		
GBC									1.00	0.70																
DBC									0.70	1.00																
CAS											1.00															
BLI												1.00	0.50	0.30	0.50	0.10	0.10	0.50	0.10	0.50	0.10					
BNS												0.50	1.00	0.50	0.30	0.10	0.30	0.30	0.05	0.30	0.05					
FLO												0.30	0.50	1.00	0.30	0.10	0.30	0.30	0.05	0.30	0.05					
BEV												0.50	0.30	0.30	1.00	0.50	0.70	0.90	0.40	0.50	0.40					
CAR												0.10	0.10	0.10	0.50	1.00	0.70	0.70	0.40	0.10	0.40					
HYB												0.10	0.30	0.30	0.70	0.70	1.00	0.80	0.40	0.10	0.40					
PEV												0.50	0.30	0.30	0.90	0.70	0.80	0.80	0.40	0.50	0.40					
FCV												0.10	0.05	0.05	0.40	0.40	0.40	0.40	0.40	1.00	0.60	0.70				
THR																				1.00	0.60	0.70				
FOR																				0.60	1.00	0.60				
MOD																				0.70	0.60	1.00				
FVR																							1.00	0.95	0.90	
FVC																							0.95	1.00	0.92	
FVU																							0.90	0.92	1.00	

Estimated cross-technology Spearman correlations for the following technologies: GTC: gasoline-substitute from biomass or a mixture of biomass and coal through thermochemical conversion pathways; JTC: jet fuel-substitute from biomass or a mixture of biomass and coal through thermochemical conversion pathways; ETC: electricity from biomass or a mixture of biomass and coal through thermochemical conversion pathways; COL: coal power without carbon capture and storage; GAS: combined cycle natural gas power without carbon capture and storage; CCS: coal power with carbon capture and storage; GCC: combined cycle natural gas power with carbon capture and storage; GBC: gasoline-substitute from biomass through a biochemical conversion pathway; DBC: diesel-substitute from biomass through a biochemical conversion pathway; CAS: compressed air energy storage; BLI: utility-scale lithium-ion-based batteries; BNS: utility-scale sodium-sulfur-based batteries; FLO: utility-scale flow batteries; BEV: light-duty battery electric vehicles; CAR: light-duty advanced internal combustion engine vehicles; HYB: light-duty hybrid vehicles; PEV: light-duty plug-in hybrid vehicles; FCV: light-duty fuel cell vehicles; THR: Gen III/III+ nuclear power; FOR: Gen IV nuclear power; MOD: small and medium factory-built nuclear power; PVR: residential photovoltaic solar power; PVC: commercial photovoltaic solar power; PVU: utility photovoltaic solar power.

# Appendix B

## Appendix to Chapter 2

### B.1 The Latent Dirichlet Allocation Model

The general process to model the data begins by preprocessing the text, constructing a document-term matrix (DTM), and fitting an unsupervised model to the data. I use the R package “tm” (Feinerer et al., 2008) to preprocess the text and create the DTM and the R package “topicmodels” (Gruen and Hornik, 2011) to handle the topic modeling.

The preprocessing step involves six iterations over the corpus. First, I remove the metadata, which was stored in the first lines of the scraped files. Next, I strip excess white space, remove capitalization, and remove punctuation. Next I delete stopwords, commonly used words which carry little substantive meaning, such as “the” or “and,” using the tm package's list of English stopwords. Finally, I stem all words in the corpus using the “Snowball” package (Hornik, 2007). Stemming removes suffixes, such that the same word used in a different part of speech is recognized as the same word. For example, with stemming, the words “position,” “positioned,” and “positions” would be reduced to just the single word “position.” While stemming could potentially introduce

additional bias into the analysis, the additional power gained by reducing the complexity of the underlying data is typically considered a worthwhile tradeoff.

In the next step, I construct the document term matrix. The DTM is a matrix with  $N$  rows and  $D$  columns, where  $N$  is the number of unique stemmed words in the corpus and  $D$  is the number of documents in the corpus. The DTM is a highly sparse matrix; in one example dataset I collected of 2,176 patents, there were 9,115 unique word stems and only 0.37% of the cells in the DTM were non-zero. The DTM is the basis for all subsequent text processing. By simplifying the data to only the DTM, two important and related features of the data are removed: 1) the ordering of words (including syntax and grammar), and 2) the proximity of words. These restrictions together are known as the “bag of words” assumption. Clearly, representing the corpus of documents by the DTM greatly simplifies analysis, but as with any simplification that removes information, important features of the raw data may be lost.

The approach I take to classifying the patent corpus is a probabilistic topic modeling approach using the latent Dirichlet allocation (LDA) model. This methodology is inspired by Blei and Lafferty (2007) and described in detail in Blei (2010). LDA requires that the number of topics be specified ex-ante (in a similar manner to other approaches I have experimented with, such as k-means, but unlike other possible methods, such as hierarchical clustering). A potential avenue for subsequent theoretical work could attempt to develop guidelines for thinking about the “optimal” (in some as-yet undefined way) number of topics to model.

The basic logic of LDA is to uncover the underlying structure of documents that likely generated the DTM extracted from the corpus. A “topic” is defined as a probability distribution over a finite vocabulary of words. Documents within the corpus have a

distribution over the topics, while each word within the document is a draw from the distribution of words conditional on a topic. Beginning to move to a full parametric model, the distribution of unique words is Dirichlet<sup>46</sup> (i.e. the probability of a word occurring in a topic has a Dirichlet distribution) and the distribution of proportions of the topics within a document is a second Dirichlet distribution. For each of the  $N$  documents, the topics within the document have a Multinomial distribution with parameter drawn from the Dirichlet distribution of topics within the document. The individual words are drawn from a second Multinomial distribution conditional on the topic (drawn from the Multinomial distribution of topics) and the distribution of unique words in a topic (Gruen and Hornik, 2011).

More formally, each of the  $K$  topics,  $\beta_1, \dots, \beta_K$  is defined by a distribution over the entire vocabulary of  $N$  words. Each of the  $D$  documents are composed of some relative frequency of the  $K$  topics, where  $\theta_{d,k}$  is the proportion of the  $k^{\text{th}}$  topic in the  $d^{\text{th}}$  document. Similarly, words within a document are assigned to topics with  $z_{d,n}$  the topic assignment ( $z = 1, \dots, K$ ) for the  $n^{\text{th}}$  word in the  $d^{\text{th}}$  document. The actual data from the DTM is  $w_{d,n}$ , the (count or binary existence) of the  $n^{\text{th}}$  word in document  $d$ . This setup frames the natural language processing statistical model as a missing data problem, hence the name *latent* Dirichlet allocation. The probability model that describes the data-generation process is:

$$p(\beta, \theta, z, w) = \prod_{i=1 \dots K} p(\beta_i) \prod_{d=1 \dots D} p(\theta_d) \left( \prod_{n=1 \dots N} p(z_{d,n} | \theta_d) p(w_{d,n} | \beta, z_{d,n}) \right) \quad (\text{B.1})$$

---

<sup>46</sup> The Dirichlet distribution is a multivariate generalization of the Beta distribution. In each dimension, Dirichlet variates are bound on  $\{0,1\}$ . Dirichlet distributions are typically used to model probabilities of (>2) rivalrous events. The canonical examples of the use of Dirichlet distributions are modeling the lengths resulting from cutting a string of length 1 into several pieces, and the ratios of colored balls in an (Polya) urn.

In words, this states that the (unconditional) distribution of topics is independent of the proportion of topics within a document which is independent of individual topic assignment of words conditional on the proportion of topics within the document, which is independent of the observed words conditional on the distribution of topics in the corpus and the assignment of topics to words in documents. The data-generation process can then be thought of as topics being generated for the entire corpus and topic frequencies generated for each document. Then, potential words within documents are assigned to topics (the same word can have positive probability in different topics), and finally observed words are generated given the overall corpus-level distribution of topics and the distribution of words within topics (Blei, 2010). Inference for the unobserved variables is made by conditioning on the observed words and using the Gibbs sampler to estimate the joint posterior distribution of the unobserved variables.

I base my implementation of the LDA algorithm on the R implementation provided in the “`topicmodels`” package (Gruen and Hornik, 2011). I fit the data with a full Bayesian model with diffuse weakly informative priors, using a Gibbs sampler. The LDA model estimates two distributions of interest: the distribution of topics within a document (for each document), and the distribution of topics within the corpus. Potential document-level covariates are the modal topic within a document (e.g. the focus of Kaplan and Vakili, 2011), or the (multi-dimensional) relative frequency of topics within a document. One potential application of the latter set of covariates could use matching algorithms on the vector of topics to identify similar pairs or groups of documents. Figure 2.5 gives a sense for the classification procedure described in this section.

# Appendix B

## Appendix to Chapter 3

### C.1 Internal Rate of Return for a Wind Farm

We determined the financial attractiveness of wind investments in different sites by calculating the project-internal rate of return (IRR). Virtually all Chinese wind projects registered under the CDM determine their additionality by comparing their project IRR with an industry benchmark (UNFCCC, 2012b). The projects used 8% rate of return on total investment and 10% on invested equity as industry benchmarks, based on guidance for power-sector investments issued by the central government (SPC, 2002).

The internal rate of return (IRR) of project  $i$  is the discount rate at which the project breaks even, i.e., at which the present value of all future cash flows in time periods  $\tau = 1 \dots T$  is equal to the initial capital investment  $Initial\_Invest_i$  in  $\tau = 0$ :

The internal rate of return (IRR) of project  $i$  is the discount rate at which the project breaks even, i.e., at which the present value of all future cash flows in time periods  $\tau = 1 \dots T$  is equal to the initial capital investment  $Initial\_Invest_i$  in  $\tau = 0$ :

$$(i) 0 = \sum_{\tau=0}^T \frac{Net\_Cash\_Flows_{i,\tau}}{(1 + IRR_i)^\tau} \quad (C.1)$$

$$= -Initial\_Invest_i + \sum_{\tau=1}^T \frac{Electricity\_Sales_{i,\tau} + CER\_Revenues_{i,\tau} - O\&M_{i,\tau} - Taxes_{i,\tau}}{(1 + IRR)^\tau} \\ + \frac{Initial\_Invest * (1 - Minimum\_Book\_Value)}{(1 + IRR)^T}$$

Chinese wind projects have two principal sources of cash inflows over the project lifetime: Electricity sales to the state-owned grid operator or, if registered with the CDM, offset credit sales:

$$Electricity\_Sales_{i,\tau} = 8760 h * Capacity\_Factor_i * Electric\_Capacity_{i,\tau} * Electricity\_Tariff_i \quad (C.2)$$

$$CER\_Revenues_{i,\tau} = 8760 h * Capacity\_Factor_i * Electric\_Capacity_{i,\tau} * Grid\_Emission\_Factor_i \\ * CER\_Price_i \quad (C.3)$$

The two main cash outflows are operation and maintenance (O&M) expenditures and taxes:

$$O\&M_{i,\tau} = Initial\_Invest_i * (Fixed\_O\&M\_Share_i + Insurance\_Premium_i) \quad (C.4)$$

$$Taxes_{i,\tau} = Income\_Tax\_Rate_i * (Electricity\_Sales_{i,\tau} + CER\_Revenues_{i,\tau} - O\&M_{i,\tau} - Depreciation_{i,\tau} \\ - Interest_{i,\tau}) + VAT_{TaxRate} * VAT_{TaxSurcharge} * Electricity\_Sales_{i,\tau} \quad (C.5)$$

In Equation (C.5),  $Depreciation_{i,\tau}$  and  $Interest_{i,\tau}$  are defined as:

$$Depreciation_{i,\tau} = \frac{Initial\_Invest * (1 - Minimum\_Book\_Value)}{Depreciation\_Time} \quad (C.6)$$

$$Interest_{i,\tau} = Interest\_Rate_{i,\tau} * Debt\_Outstanding_{i,\tau} \quad (C.7)$$

$$= Interest\_Rate_{i,\tau} * (1 - \frac{\tau}{T} * (1 - Share\_of\_Equity) * Initial\_Invest)$$

The inputs to the IRR calculation are taken from five sources (see Table C.1): First, project variables such as turbine size, number of installed turbines and project location were taken from a database of all wind farms in China installed by the end of 2012 (Huaxia, 2013). Second, power production is modeled based on high-resolution wind resource data (3TIER, 2015) and the power curve of a representative GE 1.5 SL turbine (1.5 MW rated capacity) with a rotor diameter of 77 m and a hub height of 80 m (Carrillo et al., 2013). Robustness tests were conducted using a large database of wind-turbine power curves (Carrillo et al., 2013; INL, 2006). Third, additional project-specific variables such as initial capital investment, electricity tariffs and grid-emission factors were interpolated based on data available for the 1,494 CDM projects (IGES, 2015b). Fourth, for generic financial variables that are systematically collected for all CDM projects such as the project life-time, income taxes and depreciation times we took the average over all Chinese wind CDM projects (IGES, 2015b). Fifth, for variables for which no systematic CDM data was available, including the minimum booking value of assets at the end of the life-time, O&M expenditures, insurance premiums and VAT tax surcharges, we took representative values from the CDM-PDDs.



**Table C.1: IRR Model Inputs and Assumptions**

Variable	Value	Unit	Source	Reference
Project lifetime T	21	Years	Representative value	CDM-PDDs
Initial_Invest	995-1,685	USD/kWh	Interpolated from PDD data	(IGES, 2015b)
Minimum_Book_Value	5	%	Representative value	CDM-PDDs
Capacity_Factor	0-32.9	%	Derived from wind resource	(3TIER, 2015) <sup>1</sup>
Electric_Capacity	0.25-300	MW	Wind farm database	(Huaxia, 2013)
Electricity_Tariff	379-794	RMB/MWh	Interpolated from PDD data	(IGES, 2015b)
Grid_Emission_Factor	0.724-1.152	tCO <sub>2</sub> /MWh	Interpolated from PDD data	(IGES, 2015b)
CER_Price	0/10	EUR/tCO <sub>2</sub>	Representative value	(IGES, 2015b)
Fixed_O&M_Share	2	% of Initial_Invest	Representative value	CDM-PDDs
Insurance_Premium	0.5	% of Initial_Invest	Representative value	CDM-PDDs
Income_Tax_Rate	25	%	Average value	(IGES, 2015b)
VAT_Tax_Rate	17	%	Representative value	CDM-PDDs
VAT_Tax_Surcharge	4	%	Representative value	CDM-PDDs
Depreciation_Time	15	Years	Average value	(IGES, 2015b)
Interest_Rate	8	%	Representative value	CDM-PDDs
Share_of_Equity	25	%	Representative value	CDM-PDDs

<sup>1</sup> Wind turbine power curve data was taken from (Carrillo et al., 2013; INL, 2006).

## C.2 Chinese Policies Affecting Wind Energy Development

Policy and plans in China have shaped wind power development. Through a variety of mechanisms, central government authorities, and in some cases provincial governments, have created and implemented programs affecting the rate of deployment of onshore wind energy. These policies and plans are summarized in Table C.2.

**Table C.2: Key Policies and Plans Affecting Onshore Wind Power in China 1994-2011**

<b>Year</b>	<b>Name</b>	<b>Content</b>	<b>Authority</b>
1994	Regulation on wind farm grid connection management	Wind power purchase price set at 'reasonable profit' levels Incremental cost borne by grid corporations	Ministry of Electric Power (MEP)
1995	Double Increase Programme	1.2 billion RMB of preferential loans to install 120 MW (reached ca. 25 MW); plan to develop two manufacturing bases for towers, hubs and later nacelle, braking and yaw systems	MEP
1996	New and renewable energy development outline	Target of 300 MW by end of 2000, 1,000 MW by 2010; plan to develop production capacity of 200 kW+ turbines	State Planning Commission (SPC)
1996	National High Tech R&D Program (863 Program)—Ninth Five-Year Plan	Over RMB 60 million was made available for renewable energy R&D. R&D programs included a focus on developing 600 kW wind turbines with 40 percent local content for all new wind power projects.	MOST and SDPC
1996	Loans for Wind Farm Development	This program gave priority access to reduced interest domestic loans for wind farms with rates up to 50 percent lower than current commercial rates, with a preference given to projects using locally manufactured wind turbines. It was used to develop several small demonstration wind projects.	SDPC and MOST
1997	Ride the Wind Program	400 MW by the end of 2000 Develop 300 and 600 kW turbines with 60% domestic content 23.6 million RMB to develop two 'national teams' (domestic manufacturing and S&T bases)	SPC
1998	Further Modifications to Import Duties on Wind Turbine Components	In 1998 further differentiation between domestic and imported turbine components was made when components were exempted from value-added taxes (VAT) and turbines were not, to further promote domestic wind turbine manufacturing. Customs duty regulations vary across components and are applied differently to firms with different ownership structures.	MOF
1999	Notice on next steps in developing renewable energy	Reduction of 2% on bank loans for renewable energy projects; Power purchase price set at 'production cost incl. interest rates + max 3%' for projects using foreign manufacturer, '+min 5%' for projects using domestic manufacturer.	SPC and Ministry of Science and Technology (MOST)
2000	National Debt Wind Power Programme	600 million RMB government loans to develop 73.5 MW of wind farms	State Economic and Trade Commission (SETC)
2001	10th Five Year Plan—key points for energy development	Continuation of 'ride the wind'; target to install 500MW of wind power within 10th FYP timeframe; Plan to increase localization rate from 40% to 70%	SPC
2001	10th Five Year Plan for the new and renewable energy industry	Target of 1,200 MW cumulative capacity by the end of 2005, concentrated in 100 MW+ parks; target of 150–200 MW annual domestic production capacity, focus on 600 kW+ turbines	SETC
2001	National High Tech R&D Program (863 Program)—Tenth Five-Year Plan	The continued implementation of the 863 Program as part of the Tenth Five-Year Plan included support for the development of megawatt-size wind turbines, including technologies for variable- pitch rotors and variable-speed generators, and supported the RD&D of many Chinese companies' early development of MW-scale wind turbine technology.	MOST

2001	Value-Added Tax Reductions on Wind Electricity	Reduced the VAT for wind electricity from 17 percent to 8.5 percent, resulting in a reduction in the price of wind power electricity.	MOF and the State Administration of Taxation
2003	Concession projects	Tenders for large scale (100–300 MW) wind parks, awarded based on price and localization rate; minimum localization rate set to 50% (2003) and 70% (2004–2007)	National Development and Reform Commission (NDRC)
2005	Kyoto Protocol and CDM	Chinese wind projects can earn carbon credits to be sold in global carbon markets	UNFCCC
2005	Measures for Operation and Management of Clean Development Mechanism Projects in China	Set the domestic guide-lines for CDM project development in China in line with China's role as a recipient of CDM-related carbon finance under the Kyoto Protocol for certified emissions reductions by projects that are shown to reduce GHG emissions.	State Council decision, implemented by NDRC in consultation with MOST and the MFA
2005	Renewable energy law	Requirement of compulsory grid connection and full purchase of renewable power; Electricity surcharge to cover RE cost, set at 1 RMB/MWh in 2006; 2 RMB in 2008, 4 RMB 2009 and 8 RMB since 2011	National People's Congress (NPC)
2005	Notice on the Relevant Requirements for the Administration of the Construction of Wind Farms	Clarified the project approval process and criteria for wind projects, including proximity to the grid and the percentage of domestically manufactured equipment utilized. NDRC had to approve all projects greater than 50 MW, while provincial or local DRCs could approve those less than 50 MW.	NDRC
2006	Renewable energy price and cost-sharing management	Power pricing mechanisms: (1) tendered concessions; (2) fixed provincial feed-in tariff (FIT), or (3) government approved price per project; FIT should be equal to cost difference with local desulphurized coal price	NDRC
2006	National High Tech R&D Program (863 Program)—Eleventh Five-Year Plan	The 863 Program as updated for the Eleventh Five-Year Plan supported the development of MW-size wind turbines, including technologies for variable-pitch rotors, variable-speed generators, and commercialization of wind turbines of 2 to 3 MWs in size.	MOST
2007	Medium and long term RE development plan	Target of 10% of total energy use from renewable sources by 2010; 15% by 2020; targets of 5 GW wind power by 2010; 30 GW by 2020; including 1 GW of offshore power. Mandatory market share (quota) placed on power companies with more than 5 GW of generation capacity should have 3% of non-hydro RE by 2010 and 8% by 2020.	NDRC
2007	Interim Measures on Revenue Allocation from the Renewable Surcharge	Aimed at improving equity among provinces in bearing the costs of renewable electricity, this measure established an equal-ization program requiring provincial grid companies to exchange their shortfall or surplus of surcharges with grid companies from other regions.	NDRC

2008	11th Five Year Plan for Renewable Energy	Target of 10 GW wind power by 2010, 5 parks with 1 GW+ capacity, 1 GW of offshore; target to achieve domestic batch production capabilities of 1.5 MW+ turbines, develop 3 MW+ and offshore turbine production capacity. This plan established seven wind power bases of at least 10 GW in size in Gansu, Xinjiang, Hebei, Jilin, eastern and western Inner Mongolia, and coastal Jiangsu, which together would total 138 GW by 2020.	NDRC
2008	Interim measures for administration of special funds for the wind power equipment industry	R&D fund for 1.5 MW+ turbines: 600 RMB/kW for the first 50 turbines produced (i.e., 45 million RMB for 50+ sets of 1.5 MW turbines); Blades, gearboxes and generators must come from Chinese controlled firm.	Ministry of Finance (MOF)
2009	Notice on Improving Grid-Connected Wind Power Tariff Policy	Nationally determined feed-in tariff (FIT): The program included four standardized national feed-in tariffs for wind power development with tariff levels varying by wind resource class: category I, RMB 0.51 per kWh; category II, RMB 0.54 per kWh; category III, RMB 0.58 per kWh; category IV, RMB 0.61 per kWh.	NDRC
2009	Amendment to Renewable energy law	Government system of safeguards for full purchase of RE power; Grid managers mandated to increase investment and research on grid development and its capacity to absorb RE power (smart-grid and power storage); Support from RE fund for grid connection when prohibitively expensive	NPC
2009	Notice on Abolishing the Localization Rate Requirement for Equipment Procurement in Wind Power Projects	This notice removed the long-standing local content requirement that favored domestic over imported wind technology in the Chinese market.	NDRC
2011	Notice on Strengthening the Management of Wind Power Plant Grid Integration and Operation	This notice was directed at reported problems with wind integration resulting in frequent wind power curtailments. The regulation included the establishment of new grid codes to address challenges with the interconnection of wind turbines. It also requires all wind farms to obtain NEA approval in order to receive the feed-in tariff subsidy.	NEA and the China National Standardization Commission
2011	Provisional Management Methods for Wind Power Forecasting	Aimed at improving wind power integration by better predicting when wind power will be available to the grid, required all grid-connected wind farms to install forecasting systems.	NEA