UNIVERSITY OF CALIFORNIA

Los Angeles

Three Essays in Sustainable Operations Management: Climate Change Strategies, Safety, and the Role of Information on Conservation

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Management

by

Christian Blanco

2017

ABSTRACT OF THE DISSERTATION

Three Essays in Sustainable Operations Management: Climate Change Strategies, Safety, and the Role of Information on Conservation

by

Christian Blanco

Doctor of Philosophy in Management

University of California, Los Angeles, 2017

Professor Felipe Caro, Co-Chair

Professor Charles J. Corbett, Co-Chair

In this thesis, we empirically examine different strategies that firms pursue to develop sustainable and safe operations. In the first chapter, we evaluate the profitability of carbon abatement projects that firms implement over time. We find that the average payback period and marginal abatement cost of carbon abatement opportunities do not deteriorate quickly, if at all. Firms that focus more on opportunities that are directly related to their core operations (e.g., optimizing production processes) experience more favorable payback trends compared to those that focus on opportunities that are not directly related to their core operations (e.g., lighting fixtures or building insulation). In the second chapter, we measure the impact of conducting Probabilistic Risk Assessments on safety performance at nuclear plants. We find that the adoption of risk assessments is associated with a 15% decrease in the frequency of safety-related events. We show that this is associated with a 1.6 billion USD increase in annual revenue from avoided lost production. In the third chapter, we examine the impact of automating bill payments on electricity conservation behavior. Many utility companies want to increase the convenience of paying the bill, but we find that there are unintended implications on energy conservation in doing so. We find that enrollment in automatic billing is associated with a 5% increase in energy use. This suggests that the way

we pay our bills may influence the salience of electricity costs. We find evidence that providing real-time feedback may counteract this effect.

The dissertation of Christian Blanco is approved.

Magali A. Delmas

Timothy Malloy

Felipe Caro, Committee Co-Chair

Charles J. Corbett, Committee Co-Chair

University of California, Los Angeles

2017

Contents

List of Tables

List of Figures

Acknowledgments

I am grateful to the people who guided and supported me at UCLA. They made my Ph.D. experience fruitful and enjoyable.

I want to thank my advisors Felipe Caro and Charles Corbett for their guidance and support. Both of them inspired and motivated me to pursue research. I am grateful for their training and the opportunity to work with them. Our research meetings were very exciting. I am very lucky to have them as mentors[1].

The first two chapters of this thesis is co-authored with Charles and Felipe. Chapters One and Two have been submitted to leading journals in Operations Management.

I want to thank Magali Delmas. Her enthusiasm for research is contagious, and I am grateful to have the opportunity to work with her. I enjoyed participating in her weekly seminars at the Institute of Environment and Sustainability. I admire her creativity. I am thankful for her guidance and support. Chapter three of this dissertation is co-authored with Magali, and a version of this is in preparation for submission.

I was funded by the National Science Foundation for three years. Their financial support is greatly appreciated.

I had the opportunity to interact with several other professors at UCLA, and I am grateful for their time. I want to thank Tim Malloy for agreeing to be part of my dissertation committee. I want to thank Chris Tang; his energy never ends. It was a pleasure to be his teaching assistant. There are too many professors to mention here, but I am lucky to have the opportunity to learn from them.

I have many colleagues that made my UCLA experience memorable. I thank those who kept up with my corny jokes in the office: Sandeep Rath, Paul Rebiez, Hossein Jahandideh, Thomas Vangelis, Anna Saez De Tejada Cuenca, Taylor Corcoran, Araz khodabakhshian, Prashant Chintapalli, and Ali Fattahi. I want to thank Nim Patel, Carlos Hernandez, and the rest of my colleagues at the Anderson School of Management.

It would not have been a balanced life if it werent for my friends and family who shared their time with me. Thanks to Josh, Christian, Eugene, and Gerard for being my extended Filipino family in Los Angeles. I thank Jeff, Rob, Elizabeth, Catharine, Sean, and the UCLA swimming club for motivating me to stay active. I want to thank the surfers, especially Andreas and Saki who took care of me when I was stung by a stingray! I am grateful for the Catholic community at UCLA.

I'm so glad to have family in Northern California who rooted for me all the way. I want to thank Margie and Ray Carino, Tess and Alan Miciano, their families and my cousins. I thank my parents, Victoria and Danilo, and my brothers, Kuya Joseph and Kuya Martin. They believed in me even in times when I did not. Maraming Salamat! Mahal ko kayo.

[1] In fact, I am so lucky that I was admitted twice!

Vita

Education

University of California, Los Angeles, 2017
Ph.D. Candidate in Management (Decisions, Operations & Technology Management)

Carnegie Mellon University, 2011-2012
Ph.D. Student Engineering and Public Policy

University of California, Berkeley, 2011
B.A. Applied Mathematics (Honors)
B.A. Environmental Economics and Policy

Publications

Blanco, Christian, Felipe Caro, and Charles J. Corbett. "An Inside Perspective on Carbon Disclosure." *Business Horizons* (*forthcoming*).

Blanco, Christian, Felipe Caro, and Charles J. Corbett. "The State of Supply Chain Carbon Footprinting: Analysis of CDP Disclosures by US Firms." *Journal of Cleaner Production* 135 (2016): 1189-1197.

Nelson, James, Josiah Johnston, Ana Mileva, Matthias Fripp, Ian Hoffman, Autumn Petros-Good, **Christian Blanco**, and Daniel M. Kammen. "High-resolution Modeling of the Western North American Power System Demonstrates Low-cost and Low-carbon Futures." *Energy Policy* 43 (2012): 436-447.

Ascenzi, Maria-Grazia, **Christian Blanco**, Ian Drayer, Hannah Kim, Ryan Wilson, Kelsey N. Retting, Karen M. Lyons, and George Mohler. "Effect of localization, length and orientation of chondrocytic primary cilium on murine growth plate organization." *Journal of Theoretical Biology* 285, no. 1 (2011): 147-155.

Chapter 1 Operational Response to Climate Change: Are Carbon Abatement Projects Becoming Less Profitable Over Time?

Firms around the world need to find ways to reduce their carbon footprint, preferably in ways that are profitable or cost-effective. Should firms expect carbon abatement projects to become less profitable over time as they pursue "low-hanging fruit" first? We explore this question using data collected by CDP (formerly the Carbon Disclosure Project) on over 11,000 projects implemented by 978 firms worldwide over 5 years. We find that the average payback period is lengthening by three weeks per year, but that there is substantial heterogeneity among firms in payback trends. Firms that focus more on opportunities that are directly related to core operations (e.g., optimizing production processes) experience more favorable payback trends relative to firms that pursue more opportunities that are not directly related to core operations (e.g., lighting fixtures or building insulation). We find no evidence that marginal abatement costs are deteriorating over time, even as firms implement profitable opportunities. Our results indicate the importance of exploring energy efficiency from a dynamic perspective, instead of the static view prevailing in the literature so far.

1.1 Introduction

Firms are under increasing pressure from investors, regulators, customers, and other stakeholders to limit and mitigate their environmental (Delmas and Toffel 2004) and climate change impact (Reid and Toffel 2009). Managers, members of the C-suite, investors, and policymakers use the information in climate change surveys to understand climate change risks and opportunities (Kolk et al. 2008; Reid and Toffel 2009; Chatterji et al. 2009). For instance, Blanco et al. (2017) report that at 42% of the firms they surveyed, CEOs use the climate change-related information collected as part of their CDP (formerly Carbon Disclosure Project) survey response. Simply reporting greenhouse gas (GHG) emissions is no longer enough; firms are expected to report trends and targets as well. Jira and Toffel (2013, p. 571) mention that "GHG emissions levels and trends are also among the most common environmental, health, and safety metrics reported to senior management[...]" Some 200 companies recently committed to setting science-based targets (Science Based Targets 2016a), inviting even greater scrutiny of their emissions trends. Often, firms set targets for an initial five-year period, and then set new targets depending on the firm's performance relative to its previous goal. For example, Coca-Cola Enterprises (Science Based Targets 2016b) set a 15% emissions reduction target in 2011; they comment that "When we reviewed our targets in 2015, we looked at what we had achieved and examined our carbon reduction roadmaps. As a result, we decided to stretch our absolute carbon reduction target to 50%." In 2013, Cisco reported having met goals set in 2006-2008 and announced new goals to reduce worldwide GHG emissions by 40% by 2017. Many firms report that they have been able to meet their initial targets with initiatives that are profitable in their own right. Does that mean that they will continue to be able to do so?

While there is a vast literature on the existence of profitable emissions reduction opportunities, often through energy efficiency, that literature takes a static view. Many scholarly studies focus on why profitable carbon abatement opportunities often do not get implemented (see, e.g., Jaffe and Stavins 1994 and much subsequent work). One construct that has been widely used since the 1970s to rank opportunities to reduce carbon emissions from most to least cost-effective is the marginal abatement cost (MAC) curve; well-known examples include Rubin et al. (1992) and McKinsey & Company (2009). These curves are essentially static in nature, implying that firms should start with the most profitable opportunities and keep going until they run out. If this view is true, it would have major implications for firms' response to climate change: as firms pursue low-hanging fruit first, carbon abatement opportunities will become less attractive, and setting ambitious targets will become increasingly costly. Could firms counteract such a trend by focusing their efforts on particular kinds of carbon abatement projects?

Our contribution is to provide a dynamic perspective of how carbon abatement opportunities evolve over time. Under the static view, profitable opportunities get depleted as firms pursue low hanging fruit first. An alternative view is that profitable opportunities 'grow back' as processes and circumstances change. We contrast these views, using global data from CDP (formerly the Carbon Disclosure Project), a non-profit organization that requests climate change-related information on behalf of 822 global institutional investors with combined assets of $95 trillion.

We examine over 11,000 implemented projects reported to CDP from 2010 to 2014. We use two commonly-used metrics of attractiveness of carbon abatement projects: payback period, as a measure of profitability, and marginal abatement costs (MACs), as a measure of cost-effectiveness, i.e., the net costs per ton of carbon emissions abated. (We provide a formal definition of MACs in Section 4.1.) We find that, on average, payback periods increase by three weeks (0.07 years) each year (or 3% relative to the average payback period of 2.3 years in 2010). Although this trend is statistically significant, the economic impact is modest. We find no evidence that MACs are decreasing over time, even as firms pursue profitable projects. Together, this suggests that projects may be becoming slightly less attractive over time, but that the effect is small. Furthermore, we find that firms can counteract this trend by focusing more on carbon abatement projects that are directly related to their core operations (e.g., optimization of key processes): firms that do so experience a slower decline in profitability compared to those who implement projects that are not directly related to core operations (e.g., lighting fixtures or building insulation). Our data do not allow predictions about the long-term evolution of profitability of carbon abatement projects. However, we believe that a relatively early assessment is worthwhile: if profitability declined substantially even in the limited horizon available to study, that could have major implications for the future of firms' carbon abatement opportunities and costs.

This paper is organized as follows. Section 2 is a review of related literature. We develop our theory and hypotheses in Section 3. We present the data and methods in Section 4, followed by our results in Section 5. We report robustness checks in Section 6. We provide managerial and policy insights of our results, limitations, and future work in Section 7.

1.2 Related Literature

We begin with a brief overview of related studies on energy efficiency, including some background on metrics that capture how attractive they are, and on the CDP process in general. We discuss literature specific to our hypothesis in the next section.

Reducing carbon emissions is a key component in limiting the impact of climate change (Pachauri and Allen 2014, p. 8), and energy efficiency has been identified as one of the most cost-effective measures to reduce carbon emissions (Worrell et al. 2009). Gillingham et al. (2009) provide a comprehensive review of some of the market and behavioral barriers associated with the slow adoption of energy efficiency, and they mention several areas in energy efficiency where empirical evidence is limited. These studies, and many of the studies mentioned in them, do not mention how the set of opportunities evolves over time.

Many papers in the energy efficiency (EE) literature focus on why profitable carbon abatement opportunities exist in the first place, implying that firms are not rational for not adopting them. Jaffe and Stavins (1994) refer to this as the 'energy efficiency paradox'. There is now a vast literature on this theme, and while there is some disagreement about the prevalence of profitable carbon abatement opportunities and the size of the actual savings (Allcott and Greenstone 2012), these studies focus on why individual opportunities are not adopted. They are static in nature because they focus on one-time adoption decisions. Instead of looking at what does not get adopted, we examine the profitability over time of carbon abatement projects that are implemented.

There are many studies that explore economic and noneconomic factors that influence the adoption of energy efficiency opportunities. For example, Anderson and Newell (2004) and Fleiter et al. (2012) show that projects with shorter payback periods are more likely to get adopted. Using data from the Industrial Assessment Centers (IAC), Muthulingam et al. (2013) find that noneconomic factors such as placing opportunities requiring high managerial attention early on the list can lead to lower adoption rates. All this is consistent with the assumption that companies tend to go for the low-hanging fruit first. The unexploited potential to reduce carbon emissions has been recognized as early as the 1970s (Blumstein et al. 1980). Even though firms have implemented many energy efficiency initiatives since then, the gap still exists today (Gerarden et al. 2015). Does that mean that low-hanging fruit grows back faster than firms pick it, or do these opportunities become less attractive over time as firms pick the low-hanging fruit first? That is the question we explore here.

We use payback period, the ratio of total investment divided by the annual savings, as a measure of economic attractiveness. Although payback period is far from a perfect metric for decision-making, it is used widely in the context of energy efficiency adoption in practice (Harris et al. 2000; DeCanio 1998; Anderson and Newell 2004). For instance, Harris et al. (2000) find that 80% of Australian firms they surveyed mention that payback period as the primary metric used to make decisions about carbon reduction opportunities. Anderson and Newell (2004) show that firms prefer opportunities with shorter payback.

We use marginal abatement costs (MAC) as a measure of cost-effectiveness. MAC is the net cost of abating an additional ton of carbon emissions from a single investment. A negative MAC indicates that a project is profitable, i.e., it results in a reduction in carbon emissions as well as a cost reduction. Reviewing several papers on whether technical change leads to an increase or decrease in MACs, Baker et al. (2008, p. 2800) find that "[...] the empirical basis for this aspect of technical change − how it affects marginal abatement costs − has been largely ignored in the construction of these [MAC] models." The papers that Baker et al. (2008) review are either analytical or use macro-economic data, but none use firm-level data. Our goal is not to explain the impact of technical change on MACs as Baker et al. (2008) do, but we are interested in quantifying whether MACs are increasing or decreasing based on actual projects implemented over time.

Many companies use MAC curves to rank carbon abatement opportunities. For instance, Tata Steel, a global metals and mining company, uses MAC curves to assess GHG abatement levers and to prioritize their climate change actions (CDP 2014b). Some companies develop their own MAC curves, while some rely on MAC curves developed by others. Metcash, a food distributor in Australia, mentions that their sustainability managers use the McKinsey & Company MAC curves in communicating carbon abatement activities to investment decision makers (CDP 2014b). These responses are consistent with the opportunities suggested by several MAC curves that have existed in the literature for over three decades (Meier 1982, Figure 3-1; Jackson 1991, Figure 4; and Laitner et al. 2003, Figure 2). The literature recommends that firms pursue opportunities with the most negative MAC first. The survey responses and the literature provide a good basis for using MACs to measure the attractiveness of carbon abatement opportunities.

Earlier studies have pointed out that solving environmental issues is related to concepts in Total Quality Management (TQM) such as continuous improvement (Corbett and Klassen 2006; King and Lenox 2001; Pil and Rothenberg 2003). However, very few studies exist at the intersection of energy efficiency and continuous improvement. Building upon the philosophy of continuous improvement, Aflaki et al. (2013a) develop a framework to help firms identify and implement energy efficiency opportunities. We explore whether these opportunities should be viewed as one-time or multi-period improvement decisions and whether the type of carbon abatement efforts firms focus on influences the profitability of these opportunities over time.

The emphasis of most energy efficiency papers is on technological improvements such as lighting (DeCanio 1998), air conditioners (Thollander et al. 2007), and building insulation (Fleiter et al. 2012). Although Thollander et al. (2007) look at opportunities related to production, they do not do so from a dynamic view as we do. Our contribution is to see if firms face different trends depending on the type of carbon abatement projects they pursue.

A few studies look at the CDP responses for other purposes related to production and supply chain management. Using open-ended responses from an early CDP survey, Kolk and Pinkse (2004) observe that one of the primary actions companies take to reduce carbon emissions involve changes in production processes. However, at the time of that study, CDP did not yet systematically collect the detailed information on investment, monetary savings, and carbon emissions reductions from these opportunities that we use. Using the CDP Supply Chain module, Jira and Toffel (2013) explore the factors that influence suppliers to disclose to CDP; they focus on the disclosure decision, rather than on the responses that we use. Starting in 2010, the CDP survey became more consistent and structured, allowing the more detailed quantitative analysis that we do.

1.3 Theory and Hypotheses

In this section, we formulate our hypothesis. Profitable opportunities may gradually deplete as firms pursue low hanging fruit first. Alternatively, it is possible that new profitable opportunities appear due to changes in company operations, competitive environment, technology, etc. Because there is no unambiguous theory or empirical evidence, we enumerate different reasons that may drive how opportunities evolve over time and formulate competing hypotheses.

1.3.1 How do carbon abatement projects evolve over time?

Here we outline theoretical perspectives on the attractiveness of carbon abatement projects, starting with reasons why they might become less attractive over time, then reasons for constant or improving attractiveness. We then formulate competing hypotheses using both payback and MACs as different metrics of attractiveness.

If the opportunity set is fixed and if firms pick low-hanging fruit first, then the remaining opportunities become less attractive. The assumption that the opportunity set is fixed is not uncommon in works related to search and adoption. Developing theoretical models of search theory and learning, Muth (1986) assumes that firms search from a fixed set of opportunities and select the most profitable one. He also assumes that firms pursue low-hanging fruit as soon as it is discovered. As a result, firms should experience an eventual plateau effect where no further improvement can be made. If this view also applies to carbon abatement we would expect that projects will eventually become less profitable.

Conversely, firms may continuously find profitable carbon abatement projects as technology, the production environment, and company operations evolve, i.e., if the opportunity set is dynamic. For instance, Rubin et al. (1992) estimate that (at that time) roughly five billion tons of carbon emissions could be profitably avoided annually from a variety of opportunities including vehicle efficiency, power plant upgrades, and industrial and commercial energy efficiency. McKinsey & Company published their estimates of global greenhouse gas abatement opportunities in 2009 (c.f. Figure 2 in Rubin et al. (1992) and Figure 1 in McKinsey & Company (2009)). They estimate a total of 38 GtCO2e of global emissions reduction opportunities, a much higher total than the earlier estimates of Rubin et al. (1992) when several sources of low-carbon electricity generation were not yet profitable. Many of the profitable opportunities listed by McKinsey & Company are the same as were available in 1992: industrial energy efficiency, lighting, landfill gas collection, and transportation. This suggests that many profitable opportunities "grow back".

In addition to changes in production and technology over time, firms may experience learning-by-doing. For instance, once managers learn how to reduce waste in one division, it becomes easier to replicate this in other areas (Lapré et al. 2000). As firms measure their carbon footprint more accurately and more comprehensively, they may discover other opportunities to reduce their carbon emissions (Blanco et al. 2017). Implementing a few carbon abatement projects can lead to lower search or implementation costs for future opportunities (Gillingham et al. 2009). Such benefits can also occur across firms as they publicly disclose their energy efficiency projects. In this case, the true opportunity set does not (necessarily) change over time, but the search cost to map the opportunity set decreases, so the set of opportunities accessible to firms expands.

We use two different metrics, payback period and MAC, to examine whether profitable carbon abatement opportunities decrease over time, and formulate competing hypotheses.

Hypothesis 1a: The mean payback period of implemented projects is deteriorating (lengthening) over time.

Hypothesis 1b: The mean payback period of implemented projects is not deteriorating over time.

Hypothesis 2a: The mean marginal abatement cost of implemented projects is deteriorating (increasing) over time.

Hypothesis 2b: The mean marginal abatement cost of implemented projects is not deteriorating over time.

1.3.2 Does the type of carbon abatement project matter?

After determining whether the set of implemented projects becomes more or less attractive over time, in terms of payback period or MAC, we examine whether this trend is associated with the type of opportunities a firm pursues. Opportunities to reduce carbon emissions can be broadly classified into one of two categories: those that are directly related to core company operations and those that are not. We define "core operations" as the key business processes where resources are transformed into goods and services (see Jacobs and Chase 2013). (This is not to be confused with "core competence" which refers to resources and skills that distinguish a firm in the market.) We distinguish two reasons discussed in more detail below. First, carbon abatement projects can lead to other unexpected benefits, so projects focusing on core operations are more likely to bring about other improvements in core operations, which are more valuable than improvements elsewhere. Second, changes within core production and services are themselves more substantial to a firm compared to changes to their periphery (lighting, building insulation, where they purchase electricity etc.). Therefore pursuing opportunities that are directly related to core operations can be associated with more profitable opportunities over time.

We provide a simple example on how the type of carbon abatement project may or may not lead to other unexpected benefits. A manufacturing firm can choose to put solar panels on the roof to reduce their emissions from the use of electricity. Because this is not directly related to their core operations, it less likely to lead to other benefits beyond energy savings. The same manufacturing firm can change the layout of their warehouse to reduce energy use to move goods. This can lead to other unexpected benefits such as saving time, minimizing errors, replacing older equipment, improving working conditions, and discovering other problems within their operations.

Firms that pursue opportunities that are directly related to core operations may find other unexpected benefits as a result. Using stylized models in quality control, Fine (1988) show that continuously improving production processes can lead to other unexpected benefits such as the discovery of other problems, which when addressed leads to even lower costs. Using data from two manufacturing plants, Ittner (1994) shows that the indirect benefits from improving quality, such as higher productivity or better inventory management, can be at least twice as large as the direct benefits attributed to waste reduction and rework. Using data on 614 publicly traded firms that own facilities that are required to report to the Environmental Protection Agency, King and Lenox (2002) find that firms that pursue waste prevention (as opposed to waste treatment) experience financial gains. Firms may be choosing the low-hanging fruit, but they may overlook some types of opportunities because some of its indirect benefits are hard to quantify (King and Lenox 2002).

Changes associated with core operations can be more substantial compared to those that are not. Firms often consider investments in opportunities not directly related to core operations as "non-strategic", that is, they do not help them gain a competitive edge (Sønderhousen 1993). Firms concentrate on "the survival of the firm", and they often consider lighting fixtures and improvements in buildings as peripheral to core operations (DeCanio 1993, p. 910). Firms that focus more on their core operations experience more rapid, substantial changes in these areas, and thus low hanging fruit grows back faster. For instance, the plant described in Rajaram et al. (1999) and Rajaram and Corbett (2002) was able to reduce energy use by 30% and water use by 50% at a starch production facility by reconfiguring process flows over the course of five years. The company kept adding new pipes and pumps, which led to new opportunities to re-optimize process flows. This led to savings in the order of $3 million (USD) without any capital-intensive investments. There are similar examples in the CDP surveys, such as optimizing transportation networks and improving manufacturing-related processes.

Hypothesis 3: Firms that focus more on opportunities that are directly related to core company operations experience more favorable payback trends than other firms.

Hypothesis 4: Firms that focus more on opportunities that are directly related to core company operations experience more favorable MAC trends than other firms.

1.4 Data and Methods

We describe the data collected from global firms and the methods we use to test our hypotheses.

1.4.1 Data from global firms collected by CDP

CDP has been successful in engaging the world's largest firms to disclose climate change-related information, including investments that lead to carbon emissions savings. Each year since 2003, CDP invites global firms with the largest market capitalization to report on their climate change impact. They send their information request on February 1 and firms can submit their response until June 30 of the same year. The now 17-page survey goes beyond simple carbon accounting, as it also includes questions on how firms manage their carbon emissions, what risks and opportunities firms see related to climate change, and more. Recognizing the complexity of their survey, CDP provides extensive support in answering the survey completely and accurately. Many companies contact CDP for assistance, and CDP's responses to us also reflect what they mention to firms. By 2014, 1,825 global firms publicly responded to CDP, representing all global industrial sectors; not all firms that respond provide information on their carbon abatement efforts.

There are some limitations in the CDP survey data (which we discuss more thoroughly later), such as the comparability of firm responses related to disclosed emissions. As we are only interested in within-firm trends and differences in those trends, lack of between-firm comparability of absolute emissions is less of an issue for this study.

We discuss how we determine the sample for our main analysis on payback periods and MACs. In 2010, CDP started asking detailed questions about investment in and monetary savings from implemented projects. A detailed list of the questions we use in each CDP survey from 2010 to 2014 is in the appendix. For example, we used the information provided in the survey question CC3.3 in 2014: "For those initiatives implemented in the reporting year, please provide details in the table below." The table they refer to allows the survey taker to list the following information for each project: investment cost, annual savings, annual carbon emissions reduction and the project type (chosen from a drop-down menu).

Our analyses on payback periods and marginal abatement costs use project-level data between 2010 and 2014. The payback period is computed by dividing the investment cost by the annual savings for each project reported on the survey questions. We remove payback periods that are greater than 100 years ($N = 306$) and have a Cook's statistic greater than 0.0001 ($N = 858$) as those are considered outliers. A step-by-step discussion of our filters is in the appendix. The necessary information to calculate MACs was available beginning in 2010, except in 2011 because in that year CDP did not ask carbon emissions reduction by individual project. Firms report annual carbon emissions reduction, investment costs I, and annual monetary savings S. We convert all costs and savings to USD using exchange rates published by the International Monetary Fund (http://data.imf.org, last accessed January 16, 2015). We use the average lifetime L of projects by type provided in the CDP reports. We also assume an interest rate of $r = 20\%$ such that $\delta = 1/(1 + r) = 0.83$. Interest rates of 20% and above are commonly found in the energy efficiency literature (Hassett and Metcalf 1993), so we use similar estimates. The MAC of a single opportunity is calculated as follows (Laitner et al. 2003, p. 6-110):

$$MAC = \frac{I - \sum_{t=0}^{L} \delta^t S}{emissions\ reduction^{*} L}. \tag{1.1}$$

We provide additional details on how we calculate marginal abatement costs in the appendix. Many firms do not complete the CDP survey every year; we only include firms that report at least twice during this period. We obtained an unbalanced sample of 11,941 projects across 978 firms for our payback period analysis, and 8,165 projects across 808 firms for our MAC analysis.

Table 1 summarizes the payback periods and MACs of projects implemented by calendar year. The overall average payback period is 2.76 years across all projects in all years. In 2010, the average payback period is 2.33 years and 2.68 in 2014. Firms report on an average of three to four projects per year, and if they are more likely to report their best opportunities, then this would reflect the low-hanging fruit firms pursue over time. The average MAC of implemented projects is −\$30.74 per ton CO_2e in 2010 and −\$28.80 per ton CO_2e in 2014.

Table 1.1: Summary statistics of payback period and MAC.

	Calendar year					
Statistics	2010	2011	2012	2013	2014	Overall
Payback period						
Mean	2.33	2.66	2.79	2.98	2.68	2.76
Standard deviation	2.73	3.05	3.82	4.02	3.28	3.57
Min	0.00	0.00	0.00	0.00	0.00	0
Max	12.63	15.55	20.74	21.88	16.76	21.88
Number of firms reporting	283	451	714	869	850	978
Total projects reported	859	1515	2639	3394	3534	11941
Marginal abatement cost						
Mean	-30.74	–	-28.09	-29.23	-28.53	-28.80
Standard deviation	29.36	–	43.91	50.70	41.75	44.90
Min	-97.23	–	-167.27	-200.54	-156.09	-200.54
Max	24.48	–	93.71	130.01	84.70	130.01
Number of firms reporting	206	–	575	753	740	808
Total projects reported	524	–	1994	2810	2837	8165

Notes: CDP did not ask firms to report carbon emissions reduction in 2011.

Now, we describe the different types of opportunities firms pursue as described in the CDP data. Table 2 shows the average payback period of opportunities by type. “Energy efficiency: Building services” is the most frequently reported type with 3,555 projects; these are associated with changes in the building such as more efficient lighting, heating and cooling, and other building energy management systems. “Energy efficiency: Processes” is the second most reported type with 3,324 projects, reducing emissions associated with operations or replacing production equipment. (A detailed explanation of each type is available in the appendix.) Whether a particular type of carbon abatement project is closely related to the company’s core operations depends on the specific sector that firms is active in. Firms around the world are classified into one of the Global Industry Classification Standard (GICS) Industries depending on their principal business activity. There are 66 GICS Industries that are classified into one of 10 GICS Sectors. We use the GICS Industry, the company, and project descriptions to guide us in determining whether an opportunity is directly related to a firm’s core operations or not. All three authors independently classified all 11 opportunity types for each of the 66 GICS Industries as to whether they are directly related to core company operations or not. (A detailed discussion of how we classify the opportunities is available in the appendix.)

We provide an example using the airline industry. Here, projects listed under transportation (use and fleet) are directly related to core operations because they are associated with the movement of goods and services. This includes, fuel optimization, air traffic management, and aircraft weight reduction. Emissions reductions with buildings are not directly related to airline operations. We do this classification for each of the 66×11 GICS Industry and opportunity type pairs.

Table 1.2: Summary statistics of payback and total number of projects by types of emission reduction projects.

Implement type	Mean payback period (year)	SD	Min	Max	Obs.
Behavioral change	1.01	2.56	0.00	20.00	617
Transportation: use	1.30	2.80	0.00	17.50	289
Transportation: fleet	1.96	3.19	0.00	17.67	474
Product design	2.00	2.79	0.00	12.00	140
Process emissions reductions	2.02	3.02	0.00	20.00	634
Energy efficiency: Processes	2.32	3.04	0.00	21.07	3,324
Low carbon energy purchase	2.92	4.69	0.00	20.65	127
Fugitive emissions reductions	3.11	2.76	0.00	14.03	83
Energy efficiency: Building services	3.18	3.61	0.00	21.39	3,555
Energy efficiency: Building fabric	3.96	4.24	0.00	21.88	679
Low carbon energy installation	5.53	5.01	0.00	21.58	763

Notes: CDP did not ask firms to report the type in 2010, so 859 projects do not have a type. There are 361 projects classified as "Other" and 36 projects were left blank by the survey taker. We do not use 2010 in the analysis for fraction of opportunities by type, but the trends capture the payback periods during this period.

1.4.2 Empirical Methods

For Hypothesis 1, we test whether payback periods are deteriorating or not, which involves estimating firm-level payback trends. We use the same methods to test Hypothesis 2 on MACs.

Our data on payback periods and marginal abatement costs have a hierarchical nature because these observations are for individual projects, which can vary in number from firm to firm and year to year. Although one could use (weighted) OLS, a more flexible approach is to use a mixed effects model, which allows the slope to vary by firm. For an introduction see Robinson (1991) and Speed (1991). The mixed effects models estimate the mean and standard deviation of the distribution of the random parameters in situations when repeated measures are in groups. This will provide us with a way to obtain both individual and average time trends using all the data in a single regression. The mixed effects model has two advantages over weighted OLS. First, Beck and Katz (2007) show that the root mean squared error (RMSE) of β_i is lower than that of weighted OLS, especially when the number of observations over time is low as is the case here. Second, the mixed effects model is more flexible and often performs similarly to weighted OLS even when the mixed effects model is misspecified (Beck and Katz 2007). We estimate the following equation:

$$payback\ period_{kit} = \beta_i \times year_t + \alpha_i + \eta_{kit}. \tag{1.2}$$

We estimate a random-intercept, random-slope model of equation (1.2) using Stata's *xtmixed* command. (Another conventional way to write equation (1.2) is $payback\ period_{kit} = \alpha_0 + \beta \times year_t + \alpha_i + u_i \times year_t + \eta_{kit}$ where u_i represents the departure of an individual firm from the average slope β and random intercepts α_i.) The mixed effects model provides the best linear unbiased predictor of individual slopes and intercepts (Robinson 1991). As is the case with any estimator, there are some drawbacks to this approach as we use predicted slopes, but we are not aware of any other well-established estimator that uses all the data in a single regression and obtains individual trends.

For Hypothesis 3 and 4, we explore the relationship between opportunity types that firms focus on and the trends they experience. We could estimate trends by computing firm-by-firm OLS, but there are a couple of challenges with this approach. First, some firms have very few observations that could lead to imprecise estimates of the trends. Second, this does not use all the data in a single regression model. Let $\hat{\beta}_i$ be the predicted payback period time trend for each firm i using the mixed effects model. Because most firms implement some projects that are closely related to their core operations and others that are not, we define f_i to be the fraction of projects pursued by firm i that are directly related to core operations, and $\overline{f}$ its average across all firms. We subtract the average fraction of projects that are directly related to core operations from the fraction pursued by firm i $(f_i - \overline{f})$ so that our regression intercepts will reflect the average payback period time trend. We estimate the following regression equation using OLS:

$$\hat{\beta}_i = \phi + \gamma(f_i - \overline{f}) + \xi_i, \tag{1.3}$$

where γ reflects how payback period trends are influenced by the fraction of carbon abatement projects implemented that are directly related to core operations. We can improve our estimates of equation (1.3) by adding frequency weights on the number of projects reported by each firm, putting more emphasis on firms that report a higher number of projects. We use the same method to test Hypothesis 4 on MAC trends.

All statistical analyses were done in R and Stata 11.2.

1.5 Results

1.5.1 Trends in payback periods and marginal abatement costs

We need to test whether the average payback periods and MACs are increasing or decreasing. Table 1.3 shows that payback periods are increasing by 0.072 years per year ($p = 0.015$). While this is significant, the magnitude of this effect is modest: it corresponds to a 3% deterioration per year relative to the average of 2.3 years in 2010. Put differently, payback periods are lengthening by three weeks per year. This supports Hypothesis 1a over 1b. Table 1.3 shows that the average MAC is increasing by \$0.267 per ton of CO_2e per year, but this is not statistically significant and negligible relative to the average of $-$\$30 per ton of CO_2e. Our sample size of 8,165 projects and 808 firms is large enough to capture small to moderate effect sizes, and a power analysis in the next section confirms that. Overall, the results are in favor of H2b, that marginal abatement costs do not worsen over time, rather than H2a.

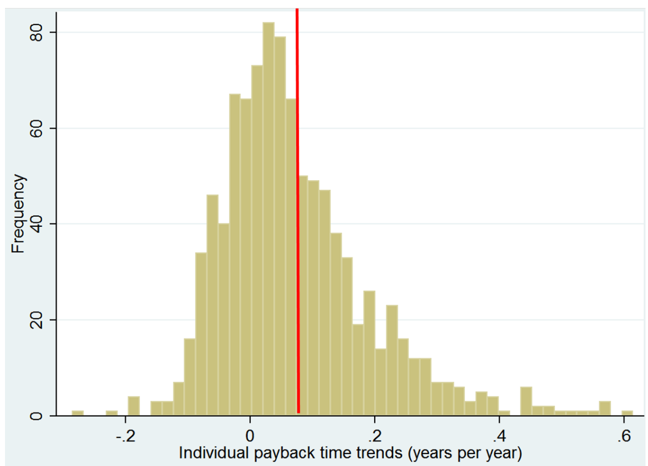
The average payback period is increasing over time, but the figure below shows that firms display substantial heterogeneity in the trends. The red line shows the average trend of 0.072 years per year. We also observe a wide variation in MAC trends, but we do not find statistical evidence that the trends are changing. We explore whether the types of project can explain part of that

Table 1.3: Payback and marginal abatement cost trends using mixed effects models.

Dependent variable:	payback period	marginal abatement cost
	(1)	(2)
Time trend	0.072^{**}	0.267
Standard error	0.030	0.491
$p-$value	0.015	0.587
Intercept	2.580^{***}	-29.887^{***}
Standard error	0.118	2.024
$p-$value	0.000	0.000
Observations (projects)	11,941	8,165
Firms	978	808

variation.

Figure 1.1: Histogram of predicted firm-level payback time trends using the mixed effects model.



1.5.2 Do trends differ for projects directly related to core operations?

Hypotheses 3 and 4 predict that firms that pursue a higher fraction of projects that are directly related to core operations experience more favorable payback and MAC trends than firms that focus less on their core operations. Table A.5 summarizes the association between core-aligned projects and payback and MAC trends. The dependent variable is the firm-level time trend $\hat{\beta}_i$, and we include a dummy variable for each GICS Sector, a coarser classification than the GICS Industries. We use the GICS Sectors instead of GICS Industries to avoid over-fitting. We identified 4,781 core-aligned and 5,565 non-core-aligned projects for the payback analysis. The key independent variable is the fraction of core-aligned opportunities constructed by dividing the total number of projects classified as core-aligned divided by the total number of projects implemented by that firm. We find that switching from 0 to 100% core-aligned projects is associated with a −0.058 years per year decrease in the average payback trend (recall that shorter payback periods are preferred). Firms can implement any fraction of core-aligned projects (as the data shows), but we find that those who focus on core-aligned projects, experience very mild deterioration in payback period compared to those that do not, supporting Hypothesis 3. The results are consistent even after controlling for GICS sectors.

Table 1.4: Impact of core operations focus on payback and MAC trends.

Dependent variable:	payback trend		MAC trend	
Frequency weighted OLS:	(1)	(2)	(3)	(4)
Fraction of opportunities directly related to core operations	-0.058^{***} (0.017)	-0.038^{**} (0.019)	0.036 (0.222)	−0.113 (0.319)
GICS Sector	–	Yes	–	Yes
Constant	0.067^{***} (0.006)	–	0.264^{**} (0.076)	–
Observations	966	966	804	804
R^2	0.023	0.071	0.000	0.019

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroscedastic-robust standard errors are in parentheses.

The fraction of projects directly related to core operations does not affect MAC trends, which in Table 1.3 are flat to begin with. We do not find support for Hypothesis 4 that firms that have a core focus face more favorable MAC trends than those that do not.

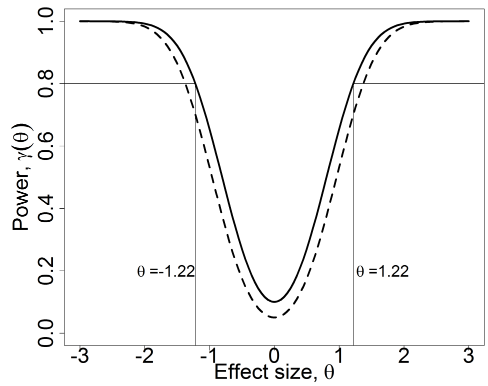
1.6 Robustness Tests

We found that payback periods are deteriorating slightly over time, and less so for firms that pursue carbon abatement projects closer to their core operations; we found no change in MACs over time. In this section, we report on several tests to assess whether our results are robust to sample size, sample selection, firms' project selection priorities, coding of projects and industry effects.

1.6.1 Potential issues with the statistical power of the tests

We found no statistically significant change in MAC over time. To assess whether this non-significant effect is genuine or due to the sample size being too small, we performed a power analysis; see Cohen (1988) and Casella and Berger (2002) for a good reference. The probability of a Type II error (not finding an effect when there is one) is equivalent to one minus the power function $\gamma(\theta)$, when the parameter of interest θ is in the region of rejection. We calculate the power of the test for the time coefficient in model (2) for the MAC analysis in Table 1.3. Figure 1 shows the power function for the mixed effects model time trend estimates at a 0.10 and 0.05 significance level. At a significance level of 0.10, the time trend estimates that can be detected with 0.80 power is $+/-$\$1.22 per ton CO2e per year for the mixed effects estimator. We consider an increase between \$1.5 to \$3 (USD) per ton of CO2e per year (or roughly 5% to 10% per year of the average MAC of implemented projects) to be modest effect sizes. See Cohen (1988, Chapter 2) for a more elaborate discussion on *small*, *medium*, or *large* effect sizes. Our estimate for the change in average MAC is \$0.267 per year, but this is not statistically significant at the 0.10 level. We consider effect sizes below 5%, i.e., below \$1.50 per year to be small or negligible. Our test can detect any effect size above 10% (or \$3 per ton of CO2e). The power analysis supports our conclusions in favor of Hypothesis 2b that MACs do not deteriorate over time.

Figure 1.2: Power analysis for MAC results in Table 1.3 model (2) at $\alpha = 0.10$ (solid curve) and $\alpha = 0.05$ (dashed curve) significance level.



1.6.2 Potential issues with the unbalanced sample

Our main analysis is based on the largest sample of firms we can construct regardless of whether firms report for all years in the study. This results in an unbalanced number of firms over time, meaning that our results could potentially be influenced by which firms happen to report in any given year. To see if our results are robust to this, we repeat our analyses with a balanced sample, the subset of firms that report every year between 2010 to 2014, yielding 124 firms for our payback trend analysis and 131 for MAC. (There are more firms in the balanced sample of MAC because we do not include 2011 when firms were not asked for total emissions abated.) The results for the balanced panel are consistent with our earlier results. We find no evidence that MACs are changing over time. Payback periods are increasing by eight weeks per year (or roughly a 7% increase relative to the average in 2010). This effect is in the same direction and about twice as large as our earlier estimate for the unbalanced sample. Note though that even the unbalanced sample only includes firms that reported in multiple periods, so the difference in magnitude cannot be attributed only to the unbalanced sample including relatively more first-time reporters.

1.6.3 Potential effect of firms' project selection policies

Although the literature provides evidence to suggest that firms prefer projects with low payback periods (Anderson and Newell 2004) and negative MACs (McKinsey & Company 2009), firms that have a dedicated budget for energy efficiency may prioritize projects differently. If a firm follows a rule that projects are implemented only if their payback period meets a specific threshold, then we will not see a deterioration over time in the payback period of implemented projects. Conversely, firms with a dedicated budget for energy efficiency may be less likely to have strict payback period thresholds when choosing a project (as investments may be seen as sunk cost). As a robustness check, we repeat our analyses separately for firms that do and do not have a dedicated budget. We use question CC3.3c in the CDP survey in 2014 (and similar questions in years before that; see the appendix for more details) to identify firms that have a dedicated budget. There are 705 firms that mention they have a dedicated budget at any point during the study. We find that firms with a budget experience an increase in average payback period of 0.095 years per year (or a 3.4% increase per year; $p < 0.01$), close to our earlier estimate of 0.072 for the entire sample. We find no evidence that payback periods are increasing for firms that do not use a budget, but this sample is much smaller. If firms that do not have a dedicated budget have strict payback period thresholds, then the results suggests that they are able to find profitable opportunities that meet these rules over time. We still find no evidence that MACs are increasing or decreasing, for firms with or without a budget. The observation that a large number of firms have a (limited) budget for energy efficiency partially explains why we only get to examine implemented projects as opposed to the entire opportunity set.

1.6.4 Potential issues related to coding project type

There can potentially be some ambiguity in determining whether a carbon abatement project is closely related to the firm's core operations or not. The results reported earlier are based on our final coding, obtained after several rounds of refinement. To determine whether our findings are robust to cases where the classification is not obvious, we repeated our analysis with several earlier versions of the coding. We performed the analysis for our first round of coding, excluding ambiguous projects, then we repeated the test excluding projects where we had a disagreement. We independently recoded the types where we disagreed after discussing why we coded the way we did. We tested the model again after agreeing that one of the authors refine the classification where we had a disagreement. Although the estimates of focusing on core operations vary from -0.037 to -0.066, they were negative and statistically significant across all model specifications. (The table of results is available in the appendix.) The robustness tests support Hypothesis 3.

We performed the same set of robustness tests for MAC, and continue to find that firms that focus more on core operations experience the same constant MAC trends as the overall sample does. We find no evidence to support Hypothesis 4.

1.6.5 Potential issues with identifying outliers

The method (i.e., Cook's statistic) we use to eliminate potential outliers may be driving the results, so we use an alternative method to identify outliers. Instead of using Cook's statistic to identify potential outliers, we implement Tukey's interquartile range test for each firm. (See the appendix for more details on Tukey's range test.) The number of projects and the sample of firms will vary based on the filters we implement. Using Tukey's interquartile range test to identify outliers resulted in 12,739 projects from 1,001 firms. We find that the average payback trend is increasing by 0.131 years per year, but this is not statistically significant ($p = 0.158$). If we implement Tukey's interquartile range test in addition to our current filters, the estimate of payback trend is 0.036 years per year, but this is not statistically significant with a $p-$value of 0.161 (with 11,069 projects and 978 firms). We repeat the same approach for MACs. If we implement Tukey's interquartile range test in place of using Cook's statistics to identify outliers, we end up with 9,168 projects and 846 firms. The time trend estimate for MACs is -0.075 with a $p-$value of 0.399. We find no statistical evidence of change if we remove outliers using Tukey's interquartile range test in addition to our current filters. This means that our results for the trends in payback and MACs in earlier sections are more conservative, meaning that profitable opportunities do not deteriorate quickly if at all.

1.6.6 Potential differences across industries in availability of opportunities

Firms in sectors with higher emissions may have more abatement opportunities that are closely related to their core operations than firms in less emissions-intensive industries. We repeat our analysis with firms that belong to sectors that account for a significant portion of total carbon emissions. Based on the 2014 CDP report, the six sectors with the largest total Scope 1 (direct emissions) and 2 (emissions from the purchase of electricity) are Utilities, Industrials, Materials, Consumer Staples, Consumer Discretionary and Energy (CDP 2014a, p. 10). Table A.9 shows that the results are consistent with our earlier conclusions. A firms that switches from none to having all core-aligned projects would experience a 0.084 decrease in its payback trends (note that a constant or negative trend suggests that profitable opportunities continue to exist). The results are consistent even after adding GICS sector controls. Models (1) and (2) support Hypothesis 3. Models (3) and (4) show that firms that pursue a higher fraction of core-aligned opportunities still see no change in MACs over time.

Table 1.5: Impact of core-aligned opportunities on payback and MAC trends for firms in high-emissions sectors.

Dependent variable:	payback trend		MAC trend	
Frequency weighted OLS:	(1)	(2)	(3)	(4)
Fraction of opportunities directly related to core operations	-0.084^{***} (0.024)	-0.070^{***} (0.024)	0.082 (0.379)	-0.178 (0.436)
GICS Sector	–	Yes	–	Yes
Constant	0.057 (0.007)	– –	0.276^{**} (0.140)	– –
Observations	601	601	515	515
R^2	0.045	0.063	0.000	0.016

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroscedastic-robust standard errors are in parentheses. The sectors with the highest emissions based on the CDP report are Utilities, Energy, Industrials, Materials, Consumer Staples, and Consumer Discretionary.

1.6.7 Potential concerns related to firm size and project scale

Larger firms may be more likely to implement projects that involve a larger investment and larger emissions reductions. For that reason, it is important that our main analysis on trends is based on payback period and MAC, which are both ratios that allow for comparisons among projects of different sizes. To check whether there is a further association between firm size and these trends, we merge firm-size information from Compustat, a database of firms and their financial information, using available identifiers (SEDOL, ISIN, Cusip, and tickers). We find a weak association between firm size, as measured by the natural log of average revenue, and payback trends (correlation is 0.03) or MAC trends (correlation is -0.01). This suggests that our results on payback and MAC trends are not confounded by firm size.

So far, we have examined whether profitability of carbon abatement projects, measured by payback periods and MACs, is deteriorating over time, but managers and investors may also want to know if they should expect a change in the number or size of opportunities they pursue. The data on the number of projects and project size is less reliable than that used earlier in this paper, so our observations on these metrics are preliminary at best. (Details are provided in the appendix.) We find evidence that firms pursue 9.6% more projects per year, and that those projects become slightly smaller over time. To measure project size, we use investment cost per project normalized by the firm-level average investment cost in order to correct for firm size and to reduce variability due to external non-controllable factors such as macroeconomic or regulation changes (see Freeman et al. (2017) for an example of this approach). We find that the normalized investment cost is decreasing by 2.4 percentage points per year ($p < 0.05$). One reason investment costs are decreasing is that firms appear to be finding more emissions reduction opportunities that require little to no additional investment (e.g. optimizing transportation load and fuel use, improving work schedules and equipment use, etc.).

1.6.8 Alternative measures of time: cumulative projects and cumulative emissions reduction

Part of our theoretical argument for why profitability of projects may not deteriorate (or even improve) over time is because firms get better at measuring emissions and finding opportunities over time. Such learning may not be caused simply by time, but by actual experience, which would be better measured using the cumulative number of projects or the cumulative emissions reductions achieved. We take the average of payback period across all projects for firm i at time t, and do the same for MACs. We estimate fixed-effects models, predicting the mean payback period and MAC as a function of the natural log of the cumulative number of projects per year, and of the natural log of cumulative emissions reduction. The results are in the appendix. We find no evidence that the cumulative number of projects is associated with changes in the average payback period or MAC. We do find evidence that a 10% increase in total cumulative emissions reduction is associated with a \$0.17 increase in average MAC ($p < 0.01$). If we suppose MACs are non-positive for a moment, this increase implies that MACs are deteriorating by 0.6% per year. This observation suggests that projects might become marginally less cost-effective over time, as firms gain more experience, but the magnitude of this impact is rather small. These analyses are consistent with our previous findings that opportunities do not deteriorate quickly.

1.7 Discussion and Limitations

In this section, we discuss interpretations, limitations, and areas for future research.

1.7.1 Discussion of Results

The purpose of this study is to find whether profitable carbon abatement projects are declining as firms implement them over time. Existing theories point to competing hypotheses, that opportunities deteriorate over time (H1a and H2a) or they do not (H1b and H2b). We find that the average payback period is increasing by three weeks per year (or a 3% increase), although statistically significant, the economic impact is modest. MACs do not deteriorate (H2b). Together, that suggests that profitable carbon abatement opportunities do not deteriorate quickly. This means that a firm's immediate past experience with profitability and cost-effectiveness of carbon abatement may be a good indicator of its future performance, at least in the short- to medium-term.

We find evidence to suggest that the trends that firms face are not completely determined by external factors alone. Firms that focus more on projects directly related to their core operations face more favorable trends. Firms that focus more on core-aligned opportunities may have two advantages. First, carbon abatement projects can generate other unexpected benefits, leading to other profitable ways to reduce emissions. Second, projects that are directly related to core operations can be more substantial. We find evidence that firms that focus more on core-aligned opportunities face more favorable payback trends compared to those that do not.

Carbon abatement efforts need not compete with the firm's attention towards core company operations. We examined firms' operational responses to climate change by examining their experience in their efforts to reduce their climate change impact. Firms that report to CDP mention that opportunities to change their operations present themselves after they assess the risks associated with climate change. The response of Walmart, a global retailer, to climate change risks includes assessing whether operational changes are needed to mitigate their adverse impact on the environment. Our study shows that mitigating climate change impacts do not necessarily have to compete with resources away from key business processes.

Managers should view energy efficiency as opportunities for continuous improvement as opposed to the one-time adoption of more efficient technologies. Many scholarly articles often focus on why specific energy efficient technologies are slow to diffuse. Although it may be the case that specific opportunities take time before all firms adopt them, when the focus is shifted to the opportunity set, we find that firms continuously find ways to reduce their carbon footprint. This means that carbon abatement efforts will require continuous attention as the production environment changes.

Firms may under-exploit opportunities to reduce their emissions in ways that are directly related to their core operations. One possible explanation for this is that the benefits associated with reducing waste are often hard to observe or difficult to measure over time (King and Lenox 2002; Klassen and Whybark 1999, p. 611). Our observation that roughly 60% of the opportunities firms report are not directly related to their core company operations suggests that managers may often overlook some of the benefits of reducing carbon emissions within their core operations. The view and the data presented here draw attention to the set of carbon abatement opportunities, and we find that they do not deteriorate quickly.

1.7.2 Limitations and future work

Despite the extensive robustness checks we performed, our work has several inevitable limitations. The quality of CDP disclosures has improved since they sent out their first survey in 2003, but there are some limitations to the information disclosed to CDP as pointed out by Kolk et al. (2008).

The sample used is taken from voluntary survey data. Reporting issues may influence some of our results. For example, firms with severely deteriorating payback periods may be less likely to report to CDP. This would imply that the average payback period may actually be longer than our estimates. However, the results of our robustness tests for the payback period for the set of firms that consistently report are very similar to those for the entire sample. If anything, the data shows the opposite of reporting bias: there are more reporting firms with deteriorating payback periods. This provides stronger evidence that our overall estimates are representative of the opportunities that firms are pursuing.

We provide some caution in extrapolating the results to conclude that all emissions can be abated profitably. It may be impossible to operate without any emissions at all, but carbon emissions may increase or decrease over time as market and production environment changes. We do find evidence that the number of projects firms report is increasing over time. This suggests that there are benefits to periodically measuring and implementing projects to reduce carbon emissions, and the reward of pursuing some of them may lead to even more opportunities in the future.

We focused our analysis on the net effect of how opportunities evolve over time, but we acknowledge that subsidies can impact the trends. Subsidies for energy efficiency may have been removed or introduced and carbon prices may have increased or decreased during the period we examine. For instance, the price of carbon on the European Union (EU) Emissions Trading Scheme (ETS) dropped from around €15 in 2010 to below €4 in 2014. A drop in the price of the EU ETS could explain some of the deterioration in payback periods for firms regulated under the ETS, but it is difficult to tie this effect to trends of individual firms based on what they report. We found very little evidence that the projects that firms report are driven by subsidies. The term “subsidy” (and variations of the term) was mentioned four times in the full set of 11,941 projects. Other related terms appeared somewhat more often but were still all very infrequent: tax credit (7x), rebates (27x), compliance (29x), and regulatory (30x). Future studies can analyze the textual content of what firms report and examine whether firms’ responses to climate change vary depending on the subsidies and regulations they face.

There are limitations with the use of MACs to rank carbon abatement opportunities. Ward (2014) mentions that choosing opportunities with the highest cost-effectiveness may not always be the optimal way to transition to a low-carbon economy, and therefore may not also be optimal for a single firm. Nevertheless, firms mention that they use MACs as a tool to identify opportunities they want to pursue starting with the most cost-effective projects.

There are other avenues for research related to targets and incentives in achieving GHG emissions. We have shown that the opportunity set is dynamic over time, but climate change related policies are often static in nature. Future studies can examine other dynamic interactions of climate change-related policies, targets, or incentives on carbon abatement efforts.

Chapter 2 Managing Safety-Related Disruptions: Evidence from US Nuclear Power

Low probability, high impact events are difficult to manage. Firms may underinvest in risk assessments for low probability, high impact events because it is not easy to quantify the direct and indirect benefits of doing so. In this paper, we measure the impact of conducting Probabilistic Risk Assessment (PRA) on preventing safety-related disruptions. We examine this using data from over 25,000 monthly event reports across 101 US nuclear reactors from 1985 to 1998. Using fixed effects models with time trends, we find that the number of safety-related disruptions reduced by 0.30 (or 15%) per month in periods after operators conducted PRA. This reduction in disruption is equivalent to the industry gaining $1.6 billion in annual revenue from avoided lost production. In robustness tests using the variation in PRA adoption across twin reactors the impact of PRA is even greater, being associated with 30% fewer disruptions. Our results suggest that even in a highly safety-conscious industry as nuclear utilities, a more formal approach to measuring risk has its benefits. We discuss areas for future research in monitoring and increasing safety compliance in supply chains.

2.1 Introduction

The direct and indirect benefits of risk assessments are difficult to measure, especially for low probability, high impact events. The direct benefits of risk assessments are hard to estimate because the counterfactual is rarely observable, and in many cases, a metric for measuring safety is not obvious. The indirect benefits of risk assessments are difficult to quantify because the link between assessments and improvements is not always clear. These are some of the reasons why managers may not invest time in conducting risk assessments.

The importance of risk management becomes salient in the case of a major accident, prompting managers to look for ways to prevent a recurrence. For instance, the 2010 Deepwater Horizon incident, the largest oil spill in the US, prompted the offshore oil drilling industry to search for new risk assessment tools. The National Commission on the BP Deepwater Horizon Oil Spill (2011, p. 251) recommended looking at the risk management experience of the nuclear industry. The operational data on nuclear power production and the 40 years of experience in Probabilistic Risk Assessment (or PRA) in this industry can provide insights on cost-effective risk management tools for the offshore oil drilling industry (Azizi 2014; Cooke et al. 2011). The purpose of PRA is to quantify the likelihood and consequences of accidents in operating complex technologies. The challenges of assessing risks in offshore oil drilling are comparable to those in the nuclear industry because both industries face low probability, high impact events in their operations (Cooke et al. 2011; National Commission on the BP Deepwater Horizon Oil Spill 2011, p. 235). Yet, the impact of PRA on safety performance is still not well-measured even in the nuclear sector (Goerlandt et al. 2016; Rae et al. 2014).

It is not clear that managers should expect any operational benefits (not directly related to safety) from implementing PRA for a couple of reasons. First, it not obvious that managers would experience any further improvements given the already strict oversight from regulators and the existing focus on safety associated with the production of nuclear power. Second, the probabilities obtained from risk assessments may not be reliable (Sornette et al. 2013), and the models used to analyze them may be incomplete (Rae et al. 2014). There is a lack of empirical tests that examine whether conducting risk assessments is cost-effective (Goerlandt et al. 2016). The slow adoption of PRA within the nuclear industry is indicative that its safety and operational benefits are not so apparent.

Examining the US nuclear industry allows us to measure the impact of risk assessments on preventing safety-related disruptions. We conducted several industry interviews to guide our data collection and validate our findings. Our interviews reveal the lack of empirical analysis of the impact of individual plant assessments on preventing unusual, safety-related events.

The contribution of this paper is to measure the impact of Probabilistic Risk Assessment on improving safety performance. We estimate this using data from over 25,000 safety-related event reports at nuclear power plants for periods before and after operators submitted their risk assessments to the regulators. We focus on the period from $1985 - 1998$ because this was when the operators first implemented PRA; it has been updated several times since then. Our estimate of the effect of PRA on the nuclear power sector does not lead to predictions for other areas. However, given the preexisting focus on safety in nuclear power, it is plausible that the effect of conducting PRA in other sectors would be at least as large as what we find in this industry.

Our results show that operators experienced 0.30 (or 15%) fewer safety-related events per month compared to an average of 2 per month in periods before they submitted their risk assessments. To put this number in perspective, we measured the association of safety-related disruptions on production: one safety-related event is associated with a 7.4% decrease in capacity factor for that month[1]. The impact of PRA on safety-related events is significant even after controlling for cumulative experience, voluntary shutdowns, and regulatory penalties (forced shutdowns). In our robustness tests, using the variation in adoption across twin reactors, the adoption of PRA is associated with a 30% decrease in monthly events, twice the magnitude compared to our base

[1]Capacity factor is the ratio of actual electricity produced divided by the maximum possible for that month. See the online appendix for the association of safety-related events on electricity generation.

results using Poisson regression. An additional robustness test using instrumental variables also shows consistently larger estimates than our base results. This suggests that potential bias in our results tends to underestimate rather than overestimate the impact of PRA on reducing the frequency of monthly safety-related disruptions.

The outline of the paper is as follows. In Section 2, we review selected related literature on disruption risk management. In Section 3, we discuss Probabilistic Risk Assessment and its history within the US nuclear power industry. We present the data in Section 4, followed by the methods and results in Section 5. We conduct robustness tests in Section 6. We discuss our results in Section 7 and provide directions for future research in Section 8.

2.2 Literature Review

In this section, we discuss selected related work. Safety is an important part of sustainable operations (Kleindorfer et al. 2005), but most of the focus has been on waste reduction (King and Lenox 2001; Lapré et al. 2000), energy efficiency (Aflaki et al. 2013b), quality (Angell 2001; Corbett and Kirsch 2001), and pollution abatement (King and Lenox 2002). Moreover, the adoption of quality and environmental management tools have been shown to improve financial performance (Corbett et al. 2005; Klassen and McLaughlin 1996), quality (Hendricks and Singhal 1997; Rothenberg et al. 2001) and compliance (Gray et al. 2015). The overlap between environmental management and improvements in operations is strong (Corbett and Klassen 2006), but the impact of risk assessments on safety and productivity is not well-measured (Sodhi et al. 2012).

It is well-established that disruptions negatively impact financial performance (Hendricks and Singhal 2005), but very little empirical work has been done to test the effectiveness of risk assessment tools on disruptions. The focus of the Operations Management literature has been mostly on mitigating disruptions in production, including models of repair and reliability. Using game theoretic models, Kim and Tomlin (2013) incorporate concepts from PRA (dependencies of component and system failures) in examining a firm's optimal investment in preventing equipment failure and in reducing the time they take to repair (recovery capacity). They find that if firms can invest in both, they may decide to invest more in recovery capacity rather than preventing failures. In some settings, repair and restoration services can be outsourced (Jain et al. 2013), but before firms can pursue these process improvements, they often need to quantify the frequency and impact of disruptions. Gupta et al. (2016) and Sodhi et al. (2012) provide a nice review on supply chain risk management. Previous studies in disruption management mostly focus on allocating resources for repair and capacity, but less on quantifying risks and the impact on safety performance.

Risk assessments and safety-related management programs can be effective in reducing risks, accidents, and deaths. Using risk assessment data from the US chemical industry from 1995 to 2000, Kleindorfer and Saad (2005) show various associations between facility characteristics, regulation, and demographics. They uncover two dimensions in managing disruptions: (1) reducing the frequency and severity of risks and (2) increasing the ability of the supply chain to absorb risks without severe consequences for operations. Levine and Toffel (2010) show that firms that adopted the ISO 9001 quality management systems standard are associated with lower death rates. We go further by quantifying not only the benefits but also comment on whether PRA is cost-effective.

2.3 Probabilistic Risk Assessment in the Nuclear Industry

What is the impact of PRA on the frequency of safety-related events? We describe the purpose of PRA, its potential benefits, and limitations. The timeline of how PRA developed in the nuclear power industry reflects some of the challenges in establishing its benefits. At the end of this section, we explain why we limit our study to $1985 - 1999$. A more elaborate historical background of PRA can be found in Keller and Modarres (2005).

2.3.1 What is PRA?

PRA is an analytical tool aimed at answering the three following questions:

1. (Accident scenario) What can go wrong?
2. (Frequency) How likely is each scenario?
3. (Consequences) What are the effects?

Risk assessment starts with identifying the hazards involved and the undesired outcome to be prevented. For instance, nuclear power operators want to reduce the chance of a nuclear reactor core melting and releasing radioactive waste into the environment. The first step of PRA is to identify events that could lead to this outcome. For example, a valve may fail to operate, preventing water from cooling the reactor core. The next step is identifying the rate of each initiating event: "how often do we expect valves to fail or pipes to break?" After estimating the frequencies, experts assess the consequences. For example, if these events happen and safety systems fail to operate, then a core meltdown can occur. The purpose of PRA is to quantify the risks of a nuclear meltdown, but in doing so, operators may discover opportunities that may also have an impact on daily operations. Operators model individual components or systems at the plant using event trees to identify sequences or a combination of failures that can lead to an accident (see LaChance et al. (2003) for more detail on PRA estimation).

We reviewed reports published by the Nuclear Regulatory Commission (NRC) and interviewed nine members of the industry to understand the variation in safety across plants and how PRA could have influenced safety performance. The members we interviewed include a former NRC chairperson, nuclear regulators, PRA practitioners, a member of the Institute of Nuclear Power Operators, a former reactor engineer, and researchers at the NRC. For example, a risk manager who implemented PRA at a nuclear plant explained to us that PRA involves a variety of engineering and statistical tools to gain insights on the weaknesses of the system, revealing opportunities for improvement. A former leading official of the NRC mentioned that the insights from PRA were used to update regulations on how long nuclear power plants can produce electricity even when one of several diesel generators is not operational. Despite several anecdotal evidence that PRA is useful, its implications for safety performance are not always straightforward (Sornette et al. 2013). Some of the challenges involved in implementing PRA are collecting data and quantifying the likelihood of events that can occur (Goerlandt et al. 2016); these estimates are subject to great uncertainty, making it difficult to prove their usefulness and cost-effectiveness for improving safety (Rae et al. 2014).

2.3.2 Historical Perspective of Probabilistic Risk Assessment in US Nuclear Power

The slow diffusion of PRA within the nuclear power industry reflects some of the uncertainty about its benefits. The earliest application of PRA at US nuclear plants was published in October 1975 (US NRC 1975). In 1979, the Three Mile Island core reactor 2 had a meltdown due to loss of cooling water. This incident led the industry and the regulators to explore more widespread use of better risk management tools. In 1983, the regulators published the first PRA procedures guide (Hickman 1983). In 1988, the regulators released Generic Letter 88-20 requiring all nuclear reactors to conduct "Individual Plant Examinations for Severe Accidents Vulnerabilities." All plant owners decided to fulfill the requirements of the individual plant examinations using PRA (US NRC 1991, p. 1-1; Poloski et al. 1999; Lochbaum 2000, p. 4). Reactors Millstone 3 and Oconee 1, 2, and 3 are the first to submit PRA in 1990; the last, in 1995, was Browns Ferry 3[2].

In the early 1980s, there was significant variation in the safety and operations at nuclear power

[2]We are grateful to Professor Mohammed Modarres at the University of Maryland for pointing us to Appendix A in NUREG 1560, the document where we obtained individual PRA adoption dates.

plants (Poloski et al. 1999; Rees 2009, pp. 100-103), but the industry was improving over time. The variation in safety and operating performance was so dramatic that this information was used at industry-wide meetings with chief executives of nuclear utilities to shame the worst performers (Rees 2009, pp. 103-105). Some of these industry-wide initiatives to improve safety include the formation of the Institute of Nuclear Power Operators (Taylor et al. 2012), changes within the Public Utilities Commission (Fremeth and Holburn 2012), and shifts in risk perception. These industry-wide trends in safety over an extended horizon create challenges for estimating the impact of PRA, but these risk assessments is a specific intervention during a well-defined window, making it possible to identify the effect of PRA.

We examine data from the US nuclear power from $1985 - 1998$ to test the impact of PRA on safety. We limit our study to the period $1985 - 1998$ for several reasons. First, this is the first time PRA was required in the nuclear sector; it has been updated several times since, but including later periods would make it harder to identify the effect of initial adoption. Second, the reporting guidelines for unusual events were standardized after 1984 (US NRC 1984); this minimizes differences in variation due to reporting. Third, we do not include events in periods after 1999 because this is the first year when some nuclear reactors were privatized (see Davis and Wolfram 2012a). Fourth, this time window gives us enough variation to compare periods before and after nuclear operators adopted PRA. Our interviews with current and former members of the industry, PRA experts, and nuclear regulators reveal contemporaneous industry-wide changes that may also contribute to improvements in safety over time, such as the introduction of the Institute of Nuclear Power Operators, changes within the Public Utilities Commission, and the role of the nuclear regulatory body. These changes contribute to improvements in safety and productivity over time making it difficult to test the impact of PRA on safety. We discuss our tests to overcome some of these challenges later.

2.4 Data on US Nuclear Event Reports and Adoption of Risk Assessment from 1985 − 1998

We collected all event reports (formally referred to as Licensee Event Reports) submitted by nuclear operators from 1985 − 1998. These reports describe an event with significant safety implications at the plant. All event reports are publicly available on the nuclear regulator's website[3]. An expert who is intimately familiar with these reports described that its primary purpose is to record operating experience at a nuclear plant to serve as a learning tool for nuclear power operators and the regulators. Event reports are used to collect data on the failure rates of various components. They are also recorded to document various events that activate safety equipment or that may prevent safety equipment from functioning properly. There are different types of reportable events, but we initially treat them as equivalent. We look at different types of events separately later. A reportable event is determined by Title 10, Code of Federal Regulations (10 CFR). Although event reports may be used to aid the regulator's oversight and compliance rules, a nuclear power regulator confirmed that they are not the basis for levying fines or penalties against reactors. The events analyzed in PRA include those described in the reports such as a reactor trip or equipment failure, but PRA is not limited to those.

Table 2.1: Variation in the year plants submitted risk assessments.

Adopted risk assessments	**Total reactors**
Year 1990	4
Year 1991	10
Year 1992	44
Year 1993	31
Year 1994	11
Year 1995	1

Operators submitted their risk assessment reports in different years (Appendix A of US NRC

[3]https://lersearch.inl.gov/LERSearchCriteria.aspx, last accessed 10-3-2016.

(1997)). Regulators required plants to conduct individual risk assessments with the intent of increasing their appreciation for risk management (US NRC 1988). Table 2.1 summarizes the number of reactors and the year they submitted their risk assessments. The earliest submission was in 1990, and the last in 1995, by which time all reactors had filed their risk assessments.

Table 2.2 shows the summary statistics. The average monthly number of events is higher in periods before plants conducted PRA, at 1.9 per month, compared to periods after submission, at 1.2 per month. The average time between two successive events is 18.5 days. The number of days between two events captures reliability because it is the time between two failures. We focus on the number of events per month, but we also explore the impact of PRA on reliability in Section 6.4.

We included monthly data on refueling and voluntary shutdowns that may influence the average monthly frequency of disruptions at the plant. An expert confirmed that reactors refuel every 12 to 18 months, and this takes anywhere from 30 to 60 days. David Lochbaum, the director of the Nuclear Safety Project, provided us with the data on when reactors refuel. Table 2.2 shows that reactors spend 7.9% of all months refueling. Two experts described that reactors perform maintenance and process improvements during refueling, so the cumulative number of times reactors have refueled may influence the frequency of monthly events. We also collected data on long-term voluntary plant shutdowns, which can take more than a year. Controlling for voluntary shutdowns is important because during these times managers perform improvements at the plant that may also influence the frequency of monthly events. There is a high opportunity cost with shutting down the plant. To put this in perspective, a 1000MW plant that operates at an average price of $80 per MWh can lose up to $60 million in revenue a month.

Regulators imposed a forced shutdown in only 1.6% of months in the sample. A former reactor engineer mentioned that these penalties are associated with significant changes in both management and process improvements at the plant. We control for this by creating a variable that captures the cumulative number of forced shutdowns over time[4].

Table 2.2: Summary statistics of reactor and monthly data.

Statistic	Mean	St. Dev.	Min	Max
Monthly events	1.614	1.731	0	16
Before risk assessment	1.961	1.903	0	16
After risk assessment	1.204	1.395	0	13
Days between events	18.452	24.244	0	200
Before risk assessment	15.369	20.069	0	200
After risk assessment	24.401	29.852	0	197
Monthly observation	Mean	St. Dev.	Min	Max
Refueling	0.079	0.270	0	1
Long-term shutdown	0.042	0.199	0	1
NRC-forced shutdown	0.016	0.127	0	1

We have additional data on reactors, such as capacity, manufacturer, and the regulatory region. We include fixed effects by plant, so controls that do not vary over time are not included here.

2.5 Methods and Results

We want to test the impact of PRA on the number of monthly safety-related events. In this section, we describe the regression model we use, then present the results on the impact of PRA on monthly events, the rate of change, and whether this impact is consistent across different types of events.

2.5.1 What is the impact of Probabilistic Risk Assessment on Safety-Related Events?

We estimate the impact of Probabilistic Risk Assessment on monthly events using a Poisson regression equation. We model the (random) number of monthly events for reactor i at time t as Y_{it} using a Poisson distribution because the number of monthly events takes discrete values from 0, 1,

[4]We are grateful to Adam R. Fremeth at the Ivey Business school at Western University for providing us historical records of forced shutdowns.

2, ... and so on. We denote the expected number of events as λ_{it}. We assume that the expected number of events has a loglinear relationship with a set of explanatory variables $\mathbf{X'_{it}}$ such that $ln\lambda_{it} = \mathbf{X'_{it}}\boldsymbol{\beta}$, where β is a set of coefficients that will be estimated. (See Greene (2012) for more details on this formulation.) The number of (observed) events y_{it} that occur for reactor i at time t conditional on reactor and time specific variables $\mathbf{X_{it}}$ is then

$$P(Y_{it} = y_{it}|\mathbf{X_{it}}) = \frac{e^{-\lambda_{it}}\lambda_{it}^{y_{it}}}{y_{it}!},\ y_{it} = 0, 1, 2, ... \tag{2.1}$$

Therefore the expected number of events y_{it} is given by $E[y_{it}|\mathbf{X_{it}}] = \lambda_{it} = e^{\mathbf{X'_{it}}\boldsymbol{\beta}}$. We estimate this equation using R software and the package *poissonmfx* for marginal effects.

We want to see whether the adoption of PRA is associated with changes in the average number of safety-related events. We create a binary variable $Adopt\ PRA_{it}$ equal to one on month t and subsequent months when reactor i submitted their PRA and zero otherwise. Because we are interested in estimating changes within a reactor, we include a dummy variable δ_i for each reactor i. We also include a linear year trend $year_t$ to control for industry-wide trends. We can include other time-varying variables, denoted by $\mathbf{X}'\eta$. We estimate the following Poisson regression equation:

$$Mean\ Events_{it} = exp(\beta_1 \times Adopt\ PRA_{it} + \beta_2 \times year_t + \mathbf{X}'\eta + \delta_{\mathbf{i}}). \tag{2.2}$$

The Poisson regression results, displayed as marginal effects, of equation 2.2 are summarized in Table 2.3. Model (1) shows the results with reactor fixed effects and a linear time trend. We find that the number of monthly events decreases by -0.30 (or 15% relative to the average before adoption) in periods after the adoption of risk assessment (p <0.01). Model (2) allows each reactor to have its own linear trend. PRA adoption is then associated with a -0.25 decrease in monthly events ($p < 0.01$). It is possible that reactors begin to experience benefits when they start conducting PRA, which takes about a year to complete (Hickman 1983, pp. 2-18 and 2-22). In model (3) in Table 2.2, *Adopted PRA* takes value 1 beginning a year before PRA submission and 0 otherwise. The results in model (3) suggests that the implementation of PRA (including one year before submission date) is associated with a −0.29 reduction in monthly events.

Table 2.3: Marginal effects of adoption on average monthly events from Poisson regression.

	Dependent variable: Monthly events		
	(1)	(2)	(3)
Adopt PRA (β_1)	-0.295^{***}	-0.251^{***}	-0.294^{***}
	(0.091)	(0.090)	(0.094)
Reactor fixed-effects (δ_i)	Yes	Yes	Yes
Time trends (β_2)	Yes	Yes	Yes
Reactor-level trends	–	Yes	Yes
Adoption is lagged for one year	–	–	Yes
Reactor-month observations	16,066	16,066	16,066
McFadden's R^2	0.07	0.10	0.10

Notes: *p<0.1; **p<0.05; ***p<0.01. In model (3), *Adopt PRA* takes 1 beginning one year before PRA is filed. Cluster-robust standard errors are provided in parentheses.

The results are consistent even after controlling for other operational and regulatory levers. We add the cumulative number of refueling months to control for experience and the cumulative number of NRC-forced shutdowns to control for regulatory pressures. The results of equation (2.2) with these controls are summarized in Table 2.4. The estimate of PRA is associated with a −0.23 decrease in monthly events after including experience on refueling ($p < 0.05$). The estimate of PRA is associated with a −0.21 and −0.23 even after controlling for voluntary and regulatory-forced shutdown. We find that The results of Table 2.4 confirm that the adoption of PRA is associated with a decrease in monthly events.

Table 2.4: Poisson regression of PRA adoption on average monthly events with refueling and shutdown controls.

	Dependent variable: Monthly events		
	(4)	(5)	(6)
Adopt PRA	-0.234^{**}	-0.219^{**}	-0.231^{**}
	(0.093)	(0.093)	(0.093)
Reactor fixed-effects	Yes	Yes	Yes
Time trends	Yes	Yes	Yes
Reactor-level trends	Yes	Yes	Yes
Cumulative Refueling months	Yes	Yes	Yes
Cumulative Long-term shutdown	–	Yes	Yes
Cumulative NRC-forced shutdown	–	–	Yes
Observations	16,066	16,066	16,066
McFadden's R^2	0.10	0.10	0.10

Notes: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors are provided in parentheses. We report the marginal effects of the results. It is possible that the estimates we have provided are still biased due to potential omitted variables or self-selection issues. If the adoption of other efforts to reduce monthly events coincides with the adoption of risk assessments across many reactors, then this omitted variable may lead to biased results. We provide robustness tests in Section 6 to address potential issues like this. The standard errors and McFadden's R^2 are the same when rounded but differ at the third decimal place.

2.5.2 Does the rate of decrease of monthly events change after submitting PRA?

Given that there is an industry-wide decreasing trend in the average number of events, we need to test whether the rate of decrease in the number of events changes after the submission of PRA. To capture this, we create the variable 'Monthly trend before submission' as follows. Suppose a reactor started operating in January 1985 and filed their risk assessment in January 1990. The monthly trend before filing takes value 1 in January 1985, 2 in February 1985, and so on, and it takes zero when they file their PRA and after. We create the variable 'Monthly trend after submission' in a similar fashion: it takes value 1 in January 1990, 2 in February 1991 and so on, and zero in all months before they submitted their risk assessment. This approach is similar to that proposed in Iyer et al. (2013) for the effect of quality management systems.

Table 2.5 shows that the monthly rate of decrease is faster in periods after operators submitted their risk assessments. Model (1) shows the rate at which operators reduced the number of events by

0.02 ($p < 0.01$) per month in periods after submitting their PRA, compared to 0.002 ($p < 0.05$) per month before submission. The conclusions are consistent after including year dummy variables in model (2). The results suggest that PRA not only reduces the number of events but also accelerates the rate at which plants improve their safety performance.

Table 2.5: Rate of decrease of events per month, before and after filing PRA, using Poisson regression.

	Dependent variable: Monthly events	
	(1)	(2)
Monthly trend before submission	-0.002^{**}	-0.001
	(0.001)	(0.001)
Monthly trend after submission	-0.016^{***}	-0.022^{***}
	(0.003)	(0.008)
Reactor fixed-effects	Yes	Yes
Year (dummies)	No	Yes
Observations	16,066	16,066

Notes: *p<0.1; **p<0.05; ***p<0.01. The marginal effects of the results are provided. Cluster-robust standard errors are provided in parentheses.

2.5.3 What is the impact of PRA on different types of events?

So far, we have treated the various types of "events" as equivalent. However, it is possible that the impact of PRA may vary depending on the type of event. The results in Table 2.6 focus on five different types of events that have the highest safety significance at the plant. These types of events can be due to either internal or external factors. Our interview with a nuclear regulator confirms that reporting of these events is standardized by the nuclear regulators and thus comparable across operators. The impact of PRA is consistent for four of these five types. The first type of event is "system actuation". These events involve an engineering safety feature to activate[5]. According to model (1), the adoption of PRA is associated with a -0.14 decrease in average monthly events. This type of event occurs on average 0.80 times per month, so the adoption of PRA is associated

[5] An example of this is when metal rods are deployed to stop any further reaction inside the core.

with an 18% decrease in the number of system actuation events per month.

The second type is "technical specifications". Operators are required to report events when they shut down due to operating under prohibited conditions. For instance, if a backup generator was declared to be out of service and not repaired within a certain time window, then operators must shut down the plant. The adoption of PRA is associated with a -0.12 (or 16%) decrease in the number of monthly "technical specification" events. The third type is "degradation", events associated with the deterioration of plant equipment such as pipes and safety barriers. The adoption of risk assessment is associated with a 27% decrease in this type of event. The fourth type is when components or systems fail or are "inoperable". Risk assessment is associated with a 10% decrease in this type of event. The fifth type is associated with events that could have prevented fulfillment of a safety function. Although the estimate is negative, this is the only type out of the five we examine where we find no evidence to suggest a significant decrease after PRA adoption. One possible reason for this is that these events are more noticeable, therefore it may be easier to prevent their recurrence without an in-depth analysis such as PRA.

Table 2.6: Poisson regression results of PRA submission on monthly events by type.

	Dependent variable: Monthly events by type				
	System actuation	Technical specification	Degradation	Inoperable	Prevent safety equip.
	(1)	(2)	(3)	(4)	(5)
Adopt PRA	-0.144^{***}	-0.123^{**}	-0.041^{***}	-0.006^{***}	-0.026
	(0.033)	(0.048)	(0.014)	(0.002)	(0.018)
Reactor fixed-effects	Yes	Yes	Yes	Yes	Yes
Time trends	Yes	Yes	Yes	Yes	Yes
Reactor-level trends	Yes	Yes	Yes	Yes	Yes
Refuel	Yes	Yes	Yes	Yes	Yes
Long-term shutdown	Yes	Yes	Yes	Yes	Yes
NRC-forced shutdown	Yes	Yes	Yes	Yes	Yes
Observations	16,066	16,066	16,066	16,066	16,066
McFadden's R^2	0.17	0.08	0.15	0.12	0.11
Mean events before PRA	0.80	0.77	0.15	0.06	0.18
Mean events after PRA	0.24	0.57	0.20	0.03	0.12
Percent reduction	18%	16%	27%	10%	14%
Note:					*p<0.1; **p<0.05; ***p<0.01

2.6 Robustness Tests

We conduct three additional tests to address potential sources of bias and one test to explore the impact of PRA using an alternative metric. First, the fact that all operators adopted by 1995 introduces the challenge of not having a control group of non-adopters. Second, there are potential issues of endogeneity: plants may adopt PRA at times when events are infrequent, or operators that are busy dealing with safety-related disruptions may be less likely to submit a PRA at that time. Third, the results may be biased due to reactors self-selecting when to submit their PRA. Fourth, we explore the impact of PRA using an alternative metric, the number of days between two successive failures.

2.6.1 Using different timing of adoption by twin reactors as control group

Most plants with two reactors submitted their PRA for both reactors at the same time, but there are a few exceptions. In this test, we exploit this variation in adoption of plants that have twin reactors with different adoption dates. One of the attractions of this test is to help deal with the lack of a suitable control group of non-adopters. Moreover, this test helps eliminate potential bias from omitted variables common to a plant that are not observable but may impact the overall safety at the plant. We estimate the impact of PRA by focusing on the difference in the number of safety-related events between twin reactors when one has adopted PRA while the other has not.

Suppose that the "true" data generating model includes an omitted variable, z_{it}, such that the number of monthly events is given by

$$events_{it} = \alpha_i + \beta_1 \times Adopt\ PRA_{it} + \gamma \times z_{it} + \beta_2 \times year_t + \mathbf{X_i'}\eta + \epsilon_{it}. \tag{2.3}$$

Examples of this omitted variable, z_{it}, could include improvements at the plant level that may vary over time but are not observable to us. If the omitted variable z_{it} is correlated with $Adopt\ PRA_{it}$, then excluding z_{it} will result in biased estimates of the association of PRA adoption with monthly events. Similarly, suppose the true data generating model for the twin reactor j is:

$$events_{jt} = \alpha_j + \beta \times Adopt\ PRA_{jt} + \gamma \times z_{jt} + \beta_2 \times year_t + \mathbf{X_j'}\eta + \epsilon_{jt}. \tag{2.4}$$

If we suppose that the omitted variable for reactors at the same plant evolves in the same way, that is $z_{it} = z_{jt}$, then we can take the difference of equations (2.3) and (2.4):

$$\Delta events_{ijt} = \alpha_{ij} + \beta_1 \times \Delta Adopt\ PRA_{ijt} + \mathbf{\Delta X_{ij}'}\eta + \xi_{ijt}, \tag{2.5}$$

where $\alpha_{ij} = (\alpha_i - \alpha_j)$, $\Delta Adopt\ PRA_{ijt} = Adopt\ PRA_{it} - Adopt\ PRA_{jt}$, $\mathbf{\Delta X'_{ij}}\eta = \mathbf{X'_i}\eta - \mathbf{X'_j}\eta$, and $\xi_{ijt} = \epsilon_{it} - \epsilon_{jt}$.

We estimate equation (2.5) for twin-reactor plants (a total of ten reactors) with different file dates. These plants are Arkansas Nuclear, Beaver Valley, Indian Point, Millstone and Nine Mile Point. There are several reasons why reactors at the same plant can have different PRA file dates. In the case of Indian Point 2 and 3, two separate operators manage them: Consolidated Edison and New York Power Authority. In the case of Arkansas Nuclear and Millstone plants, the reactors at these plant have different manufacturers. Arkansas Nuclear 1 is manufactured by Babcock & Wilcox and Arkansas Nuclear 2 is manufactured by Combustion Engineering (CE). The Millstone 2 reactor is manufactured by CE and Millstone 3 is manufactured by Westinghouse.

The test using twin reactors yield the same conclusion as earlier results, but the magnitude is twice as large. Model (1) in Table 2.7 shows the regression results of equation (2.5). The submission of PRA is associated with roughly a -0.79 ($p-$value< 0.01) difference in average events when compared to a twin reactor that has not yet adopted PRA. Models (2) through (4) include other time-varying controls, and we see that PRA adoption is still negative and statistically significant. In model (4), where we include all three time-varying controls, we find that reactors that adopted PRA first experienced a -0.68 decrease in monthly events compared to their twin reactor in periods when they have not yet submitted PRA. This is roughly a 31% decrease in average monthly events since the average count of events before PRA submission of first adopters among twin reactors is 2.2 per month. This suggests that for these reactors when we account for the possibility of omitted variable bias at the plant level, the impact of PRA is even larger compared to our earlier estimates. However, this test does not correct for potential endogeneity at the reactor level within the same plant.

Table 2.7: OLS regression results of equation (2.5) using data on twin reactors.

	Dependent variable: Δ *events*$_{ijt}$			
	(1)	(2)	(3)	(4)
$\Delta Adopted\ PRA_{ijt}$	-0.789^{***}	-0.766^{***}	-0.676^{**}	-0.683^{**}
	(0.268)	(0.273)	(0.274)	(0.276)
Plant fixed-effects	Yes	Yes	Yes	Yes
Δ Refueling months†	–	Yes	Yes	Yes
Δ Long-term shutdown†	–	–	Yes	Yes
Δ NRC-forced shutdown†	–	–	–	Yes
Observations	753	753	753	753
R^2	0.047	0.047	0.066	0.066

Notes: *p<0.1; **p<0.05; ***p<0.01. † Difference in cumulative values. Heteroskedastic-robust standard errors in parentheses.

2.6.2 Addressing potential endogeneity

An instrumental variables approach is one way to address potential bias due to endogenous variables (Angrist et al. 1996). A valid instrument is one that is highly correlated with the adoption of PRA, but that does not directly influence the number of safety-related disruptions. We construct an instrument as the fraction of other reactors that have adopted PRA within a regulatory region. A nuclear power regulator explained that each region is overseen by resident inspectors who rotate at different plants within a region. We observe that operators tended to adopt PRA around the same time as other operators within the same regulatory region. Some reports document region-specific variation among inspectors (Boss et al. 1997), and we verified this with an active member of the nuclear power community. This is a promising instrument for the adoption of PRA because it is highly correlated with individual reactor-specific adoption dates, but there is little reason to believe that adoption of PRA at other plants directly impacts the frequency of safety-related events at the focal reactors. (See the online companion for a more thorough discussion of how we construct the instrumental variables and the results.)

Using the instrumental variable, the adoption of PRA is associated with a -0.44 decrease in monthly events ($p < 0.01$). Overall our results are consistent with our conclusion in the previous section. Our robustness tests show that, if anything, the potential bias might underestimate the impact of PRA on the frequency of monthly events.

2.6.3 The impact of PRA on the best and worst performing plants

The adoption of PRA may have little to no impact on plants that already had a strong history of safety performance because they have little room for improvement. If this is the case, then our earlier approach may underestimate the impact of PRA on less well-performing plants. We test equation (2.2) by quartile of the average number of events recorded in 1988, capturing individual operators' past safety performance. Operators in the first quartile, those with the lowest number of events in 1988, are the top performers, and those in quartile 4 the worst. The results in Table 2.8 show that the impact of PRA for those in quartiles 3 and 4 is larger compared to our earlier estimates. Conversely, the top quartile did not experience a significant improvement; this may be in part because some of these plants had adopted PRA before it became a requirement.

Table 2.8: Poisson regression results (marginal effects) of PRA on monthly events by quartile of average monthly events in 1988.

	Dependent variable: Monthly events			
	Quartile 1†	Quartile 2	Quartile 3	Quartile 4
	(1)	(2)	(3)	(4)
Adopted PRA	0.030	-0.232^{*}	-0.492^{**}	-0.371^{*}
	(0.145)	(0.136)	(0.249)	(0.207)
Reactor fxied-effects	Yes	Yes	Yes	Yes
Time trends	Yes	Yes	Yes	Yes
Reactor-level trends	Yes	Yes	Yes	Yes
Observations	3,988	4,382	3,236	3,919
Mean number of events before adoption	1.23	1.93	2.28	2.54
Mean number of events after adoption	1.15	1.19	1.26	1.24
Percent reduction	NA	12%	21%	15%
Number of reactors	24	27	20	25

Notes: *p<0.1; **p<0.05; ***p<0.01. † Reactors with the lowest average number of monthly events in 1988, two years before any reactors filed their risk assessments to the regulatory body. The total number of reactors does not add up to 101 because some plants were not yet in commercial operation in 1988, therefore they did not have any monthly event data at that time.

2.6.4 What is the impact of Probabilistic Risk Assessment on Days Between Events?

So far, we used the number of monthly events to capture safety performance at the plant, but we can also examine the effect of PRA by looking at the mean time between failures. This metric captures the increase in reliability given that nuclear power plants are base load generators, that is, they produce electricity continuously[6]. Our analysis in the online companion shows that the occurrence of safety-related events is associated with a 7.4% decrease in electricity production for that month (sometimes referred to as capacity factor). This means that plants are more productive if the time between successive events is longer.

We estimate the impact of adopting PRA on the number of days between two successive events k and $k-1$ for reactor i in time t, $Days_{k,k-1,i,t}$. The unit of observation is now an individual event.

[6] See https://www.eia.gov/todayinenergy/detail.php?id=30972 for a short description of nuclear power production; last visited May 1, 2017.

We estimate the following regression:

$$Days_{k,k-1,i,t} = \alpha + \beta_1 \times Adopted\ PRA_{it} + \beta_2 \times year_t + \delta_{\mathbf{i}} + \mathbf{X}'\eta + \epsilon_{it}. \quad (2.6)$$

The various models in Table 2.9 all provide statistical evidence of increased reliability from the adoption of PRA. Model (5) shows that the average number of days between two events increased by roughly 3.87 days (or 25%) in periods after plants submitted their PRA relative to a baseline of 15.4 days between events prior to PRA.

Table 2.9: OLS regression results of the impact of PRA on the number of days between successive events.

	Dependent variable: days between successive disruptions				
	(1)	(2)	(3)	(4)	(5)
Adopted PRA	4.313^{***}	4.248^{***}	3.912^{***}	3.812^{***}	3.866^{***}
	(0.617)	(0.637)	(0.633)	(0.635)	(0.641)
Reactor fixed-effects	Yes	Yes	Yes	Yes	Yes
Time trends	Yes	Yes	Yes	Yes	Yes
Reactor-level trends	Yes	Yes	Yes	Yes	Yes
Cumulative Refueling months	–	Yes	Yes	Yes	Yes
Cumulative Long-term shutdown	–	–	–	Yes	Yes
Cumulative NRC shutdown	–	–	–	–	Yes
Observations	25,872	25,872	25,872	25,872	25,872
R^2	0.086	0.125	0.126	0.126	0.126

Notes: *p<0.1; **p<0.05; ***p<0.01. We removed observations where the number of days between two successive events is greater than 200 as we considered those potential outliers (0.01% of the data). An outlier can occur if an event was not included in the data for whatever reason, making the number of days between two events greater than it should be. This is more likely to make our results more conservative because 15 of these instances occurred in periods before PRA adoption and 40 in periods after.

2.7 Discussion

We wanted to measure the impact of Probabilistic Risk Assessment on the frequency of safety-related events at nuclear power plants. Although PRA has been around for more than 30 years, very few studies exist on the size of its impact on improving safety and operations. Our interviews with members of the industry and the Nuclear Regulatory Commission (NRC) guided us in collecting field data to conduct our tests. Our results show that the adoption of PRA is associated with a 15% decrease in the frequency of monthly events. We estimate that this reduction is equivalent to the industry avoiding $1.6 billion lost due to production disruptions.

PRA is effective even within an industry as tightly regulated and safety-conscious as the nuclear sector. We find that the impact of PRA is significant even after controlling for experience, voluntary shutdowns, and penalties (NRC-forced shutdowns). Industry-wide and reactor-specific trends over time create several challenges in estimating the impact of PRA. We included time trends, as well as tested models with reactor-specific trends, to control for potential confounding factors. Although we cannot fully control for all industry-wide improvements, our results and robustness tests remain consistent even after controlling for different time trends and testing alternative estimators. Our study may not predict the impact of PRA in other industries, but if anything we speculate that its impact may be larger for other industries where safety is not as highly regulated as with nuclear power production.

The 20-year gap between the development of PRA and when all operators adopted it is indicative that the impact of PRA on safety and operations was not clear. Although the primary purpose of PRA is to quantify the probability of a nuclear reactor meltdown, we found that the process of identifying and quantifying these risks lead to other benefits beyond obtaining these estimates. Our study does not provide any insights on the quality of these estimates, but some reports document the uncertainty and the variation in the frequency of core damage (US NRC 1997, Appendix B). We do find a decrease in the frequency of safety-related disruptions in periods during and shortly after conducting PRA. The findings suggest that even though there could be limitations to the precision of estimating risks (Rae et al. 2014), the process of measuring it has its benefits. Moreover, we find that PRA adoption is associated with faster rates of reducing the number of safety-related events compared to periods before adoption.

PRA may have indirectly contributed to standardized safety culture and performance within the industry. We observed a wide variation in safety performance before the adoption of PRA. Table 2.8 shows that the plants with the least number of safety-related events had about one less event per month compared to the worst performers. In periods after adoption, the average number of safety-related events for the best and worst performing plants were about the same. The process of quantifying risks may have unexpected benefits that may be difficult to identify beforehand.

We explored a related metric, the mean time between failures, to see whether PRA has any impact on reliability. We find that PRA is associated with roughly a four-day increase in the average occurrence between two safety-related events. Given that these events are associated with a decrease in electricity production, this increase in time between failures implies higher reliability and productivity at the plant. Although the primary purpose of PRA is to quantify the risks associated with operating complex technologies, our results suggest that it may also have indirect benefits to productivity. We find that plants with fewer safety-related disruptions are likely to be more productive and reliable compared to plants with more safety-related issues.

Our estimates suggest that conducting PRA is cost-effective. A report by the US General Accounting Office estimates that the cost of conducting individual plant PRA is between $460,000 to $1.4 million USD in today's value (or $200,000 to $600,000 in 1985 according to US General Accounting Office (1985)). The reduction in the frequency of safety-related disruptions is associated with a $13.1 million increase in annual revenue from avoided lost production. To estimate this, we merged data[7] on monthly capacity factor, the ratio of actual electricity produced divided by the maximum possible for that month, to our data set. We then estimated the association of the number of safety-related events on the capacity factor. We found that one event is associated with a 7.4% decrease in the capacity factor for that month (see the online appendix for the regression

[7]This data set is from Davis and Wolfram (2012a).

results). This means that a 0.30 reduction in the number of safety-related events is an additional 16.4 hours (2.2% increase) of monthly electricity production. This increase is equivalent to 16,400 MWh per month for a 1000MW plant. At an average electricity price of $80 per MWh, nuclear power operators can gain up to $13.1 million per year from avoided safety-related disruptions. Overall, the direct and indirect benefits of PRA may outweigh the cost to implement them.

2.8 Limitations, Conclusions and Future Research

There are limitations to our approach. The lack of a randomized controlled experiment makes it difficult to assess the true causal impact of risk assessments on reducing the frequency of monthly disruptions. This can lead to a bias in our estimates in periods after the adoption of risk assessments. Conducting randomized field experiments is costly, but our study may serve as a benchmark for future field studies in the area of risk assessments. Our robustness tests point that potential bias was likely to underestimate the impact of risk assessments on safety instead of overstating them.

Another possible limitation is that the results may not necessarily be generalizable using other types of risk assessments or in other industries. There are different types of risk assessments, but other risk assessment tools share many common features with PRA. For example, the need to collect data and identify points of vulnerabilities is a common feature in many risk assessment tools. There are other industries, such as the National Aeronautics and Space Administration (NASA), that use PRA. It may be the case that the magnitude of our estimates will vary depending on the industry, but given that the nuclear power is heavily regulated, our results on the impact of PRA could be underestimated compared to its impact in other industries with less regulatory oversight.

We offer directions on future work. Safety is an important dimension of sustainable operations, but not a lot of research looks at how safety standards extend to the supply chain. There is evidence that quality can spill over to suppliers (Muthulingam and Agrawal 2016), but it remains unclear if this applies to safety standards. Anecdotal evidence based on our interviews suggest that the nuclear industry is an example where risk assessment practices first started with the operators and were later adopted by nuclear reactor manufacturers. Future studies can examine this diffusion. Other studies have identified some of the barriers to increasing safety, including the perception that it can hinder production (Pagell et al. 2014), but we show that increasing safety performance can have indirect benefits to reliability and production. Our study is specific to the nuclear industry, but future work can examine whether the magnitude of the relationship between safety and productivity may vary by industry and determine the factors that influence those links.

The event reports can be examined in greater detail to see how operators identified various root causes of safety-related disruptions, and how this evolved over time. For example, some suggest that identifying the root causes is a critical step in avoiding the capability trap (Repenning and Sterman 2001), but troubleshooting without identifying root causes remain prevalent in practice. Earlier periods in the nuclear power industry may reflect some of these issues, and the event reports can be examined to gain insights on how managers deal with disruptions and its root causes.

We showed that quantifying risks using PRA could be a cost-effective tool to manage risks and improve safety. It is well-established that measurement does lead to improvements in quality and productivity. We show that this may also be applicable in other dimensions of risk and safety.

Chapter 3 Electronic Billing, Smart Meters, and the Salience of Energy Use

Electronic and automated payments reduce service cost and increase convenience, but do these types of payments have any implications for energy use? This question is important because electric utilities are encouraging users to adopt digital and automated ways to pay their bill. We explore this question using over 760,000 monthly billing data from over 20,000 households from the largest electric utility company in the US. We find evidence that paying bills online or through automatic payment services are associated with a 0.8% and 4.6% increase in energy use as compared to paying by check. This suggests that the convenience associated with electronic, automated methods of payment may reduce the salience of energy cost and consumption. We find that providing daily access to information through smart meters can have the opposite effect. Users who pay automatically decrease their average energy use by 4.2% in periods when they have access to smart meters. Although the magnitude may look small, these estimates are substantial when multiplied across millions of households. We find that the impact of smart meters vary by household demographics. This implies that providing daily access to information may not be sufficient for conservation behavior, and that households respond to this technology differently. Our findings indicate that transitioning to digital service operations have implications on conservation behavior that may have been previously ignored.

3.1 Introduction

Many companies have chosen to interact electronically with their customers and electric utilities are no exception. Many households have several convenient options to pay their utility bills. They can pay their bill online or enroll in automated payment services. Automated payments can be particularly convenient for customers who are worried about missing payment deadlines. Yet, very little is known whether there are unintended behavioral implications from increasing convenience through digitizing and automating payment services. In this study, we explore whether different payment methods may make electricity cost more or less salient and whether providing access to daily cost and consumption information have a similar or competing effect.

Some estimate that the benefits of digitizing business processes can reduce service costs up to 90% (a McKinsey & Company report by Markovitch and Willmott (2014)), but little work has been done to examine the conditions when digital services may or may not increase the salience of cost and consumption information. Processing electronic payments cost significantly less compared to the cost of processing payments sent by mail (Infotrends 2015, p. 14), and switching from paper billing to electronic payments can reach as much as $26.4 million in annual cost reduction[1]. Yet, the impact of increasing convenience through digital services on consumption is not well-established.

Our results show that how we process and access energy cost information may influence energy conservation behavior. We find evidence that electronic payments and enrollment in automated ways to pay are associated with an increase in energy use. This suggests that the convenience associated with these payment types may make energy costs less salient. We find that it is possible to mitigate this effect by providing households with daily electricity cost and consumption. However, this impact may vary by income and how users pay their bill. For instance, the ability to translate

[1]A paper bill cost roughly 50 cents to process, but electronic payments cost only 10 cents. The savings of 40 cents per payment is multiplied by 12 months across 5.5 million customers.

daily cost and consumption information into energy savings may vary by income, and users who prefer convenient, faster ways to pay the bill may pay less attention to real-time feedback. This has implications for targeting and engaging users in order to maximize the benefits from these digital services. Access to smart meters is a promising technology that may be necessary to induce conservation, but it may not be sufficient to induce conservation behavior across the entire population. Further research is required to explore complementary conservation strategies using smart meters.

The remainder of the paper is outlined as follows. In section 2, we discuss the relevant literature and our hypotheses on payment methods and electricity use. In section 3, we introduce the data. We provide results in section 4 and several robustness tests in section 5. We offer our policy and managerial insights in section 6.

3.2 Background and Related Literature

The period when we receive information on our energy consumption and how we pay our utility bills may influence our energy use. Using daily data of electricity consumption, Gilbert and Zivin (2014) show that users decrease their average consumption during the first seven days of when the bill arrives. They attribute this decrease to the salience of reading and paying the utility bill. Looking at monthly billing payments, Sexton (2015) show that households who switch to automatic billing payment increase their consumption compared to periods before they did so. Our paper is different in two ways. First, we explore whether electronic (not necessarily automated) ways to pay may have the same effect as automation. Second, we test whether providing access to daily cost and consumption information may or may not mitigate the impact of automatic billing.

Many households now have access to real time electricity consumption, but the evidence how users respond to this technology is mixed (Delmas et al. 2013). Although there is evidence that communication strategies can be effective in encouraging conservation behavior (Asensio and Delmas 2015), little is known how the impact of smart meters may vary by demographics. Competing arguments can be made that consumers may not always process information perfectly and translate them into savings effectively (Hanna et al. 2014). In this paper, we examine the interaction of switching payment methods and the introduction of access to real-time feedback. The interaction of real-time feedback with how users pay and household demographics is important because of its implications on conservation behavior. This is not explored in Sexton (2015) or in the papers reviewed in Delmas et al. (2013). If the impact of having access to smart meters varies across the population, then it may only benefit specific demographics. We explore whether the ability to translate information into savings are consistent with low- and high-income households. In particular, we test whether low-income households who participate in reduce rates increase or decrease energy use when given access to real-time feedback.

There are related studies that look at the role of information on conservation behavior. Jessoe and Rapson (2014) find evidence that providing high-frequency information on energy use is related to energy conservation. One explanation is that with high frequency energy use information, households are able to identify which appliances use the most electricity, and thus identify opportunities to reduce energy consumption. This information is often accessible through a personalized website provided by the utility company. The personalized website can help consumers determine the price of electricity, which may be difficult to track without a smart meter (Shin 1985; Ito 2014). In this study, we explore the implications of increasing the convenience of paying the utility bill through automation or web-based services, and whether the impact of providing access to real-time feedback may vary by household characteristics.

3.3 Hypotheses

Our recollection and perception of energy cost and consumption may depend on the granularity (e.g., monthly versus daily) of this information and how we process this when we pay our bills. We draw from behavioral and neuroscience studies to motivate our hypotheses on how electricity cost and consumption can become more or less salient depending on how we pay. Then we provide reasons why the interaction of different ways to pay and real-time feedback may or may not impact energy use.

The type of payment method may influence how consumers value goods or what they may recall about a product, its cost or the quantity. Conducting experiments of selling sports tickets, Prelec and Simester (2001) find that participants are willing to pay more when told they had to pay using a credit card compared to those who were told they had to pay by cash. Conducting three experiments to examine the impact of payment transparency on consumption, Soman (2003) find that payments types where the monetary value is more transparent, thus more salient (see Table 1, page 175 in Soman (2003)), is associated with lower consumption. The monetary value of a payment type is more transparent if the user can infer the exact value that it carries, e.g., the value of cash is more easily quantifiable than prepaid cards or credit cards. These studies suggests that different payment methods may influence the salience of certain attributes, making consumers evaluate the same good differently.

Writing has been shown to better influence recall than other methods of inscription such as typing on a keyboard. Conducting a randomized-control experiment, Smoker et al. (2009) find evidence that those who were assigned to write words from a list had a higher recall compared to those who were assigned to type it. One possible explanation for this is that writing requires more movements and time compared to typing (Smoker et al. 2009). Mangen et al. (2015) conducted a similar experiment to Smoker and colleagues (2009), but tested within subject variation on the impact of longhand and typing on word recall. They find evidence that participants who were asked to use longhand had better recall than those who typed their notes. Using Magnetic Resonance Imaging (MRI) scans on pre-literate children, James and Engelhardt (2012) find evidence that handwriting shapes and letters activated a "reading circuit" in the brain known to underlie successful reading skills but typing or tracing did not[2]. Although the studies we mentioned are not about writing a check, we use them to motivate our arguments why writing checks may have a different impact on how much users remember about their energy bill compared to using a computer to pay your bill.

Users may spend more time processing the information when paying by check compared to paying automatically, thus making energy cost less salient. The reason for this is that utility bills are traditionally mailed to households, and users pay their bill by writing the amount due and mailing a check. We expect that bills paid by check are more coupled to consumption compared to transferring funds electronically or through automated, recurring payments.

The mental accounting literature suggests that when payments and consumption are weakly coupled, i.e. payment does not evoke strong thoughts about consumption, users may allocate consumption expenditure less efficiently. Using stylized mental accounting models, Prelec and Loewenstein (1998) show that the method of payment is one of the most important determinants of coupling. Their models suggest the temporal relationship across these different payment methods influence how consumers may have different valuations when paying for the same item. For instance, they argue that paying by cash creates tight coupling because it is noticeable what is being paid for and when payment is occurring.

The convenience of paying electronically or automatically may lead to weaker coupling of cost

[2] We acknowledge that this study is about cognitive development in children and does not suggest anything about adults, but this study suggests that writing activates certain part of the brain related to reading whereas typing did not.

and consumption. Online payments are designed to make payments more convenient, reducing the time it takes the consumer to process their bill. This reduction in time that users spend in paying their bill may make energy cost less salient. One of the reasons why households do not want to adopt electronic billing is because they are worried they may miss a payment (Infotrends 2015, p. 20). This is consistent with our hypothesis that online ways to pay may make energy costs less salient. For example, the utility company in this study does not require users to login to a personalized website to pay their bill. When costs are less salient, customers may increase their consumption relative to what they typically purchase (Chetty et al. 2009; Sexton 2015). We hypothesize that switching to online or automated payments may lead to higher energy use (paying online is not automatic).

Hypothesis 1: Switching to online and automated methods of payment is associated with an increase in energy use.

Providing access to daily cost and consumption may increase the coupling of energy consumption and its cost. Traditionally, electricity consumption and information on its cost is not strongly coupled. For example, electricity is consumed on a daily basis, but bills are paid once a month (or less frequently than that in some cases). Smart meters allow customers to access information on the current marginal price of electricity[3]. Access to this type of information in real-time may make energy costs more salient.

The timing of when users pay and when they access information (if they do at all) may vary by payment type. There are studies that show how users may perceive cost and benefits differently when payments are separated from consumption. Using survey studies, Shafir and Thaler (2006)

[3]See Reiss and White (2005) for an introduction on tiered electricity pricing schedules

find that consumers may perceive the cost (or value) of goods differently when the timing of purchase and its consumption are not the same. They find that consumers may consider past purchases as free if they consume these goods or services later. Conducting lab experiments on attending sports events, Gourville and Soman (1998) find evidence that the timing of payment may influence the consumption of services. Their results show that smaller differences in time between consumption and payment leads to higher likelihood of consuming the services they initially purchased. Although these studies do not discuss the implications of smart meters, we draw from them to hypothesize that providing real-time information about consumption costs may bridge the temporal gap between when consumption and payments occur, therefore leading to changes in energy use.

We hypothesize that how users pay their bill and how they access their information are associated with changes in energy use. Automated payments may make energy bills less salient, but access to real-time feedback may counteract this effect. Some utilities now provide households with a personalized website where they can access their own data. This may lead to a stronger coupling of cost and consumption.

Hypothesis 2: The interaction of electronic and automated payments and access to real-time electricity information is associated with a decrease in energy use.

3.4 Data

3.4.1 Sample households

We test our hypotheses using electricity consumption data from Pacific Gas & Electric (PG&E) obtained through The Wharton Customer Analytics Institute (WCAI) at the University of Pennsylvania. PG&E is the largest utility provider in the US in revenue ($12.25 billion). PG&E provides service to more than 5.4 million households in California (Pacific Gas & Electric Company 2016). The company covers a 70,000-square mile service area in northern and central California. According to those who collected the data from PG&E, the data they provided is a representative random sample of the utility service area across three geographic regions in central and northern California: Central Valley, Inland Hills, and the Coast.

The data contains 1,064,857 monthly electricity bills. We describe how we select the sample of bills in our study. We removed monthly bills below 20 kWh and above 5000 kWh. We considered monthly bills below 20 kWh and above 5000 kWh as outliers or erroneous[4]. We limited our sample to households that move dwellings at most once. This limits potential confounds due to multiple relocations. After applying these two filters, we were left with 1,007,177 monthly bills. Then we removed bills that could not be merged with census data. This left us with 998,162 monthly bills. We subsequently removed monthly bills that could not be merged with the premise and the payment dataset provided by PG&E. We also excluded monthly bills outside our study period (January 2008 to December 2011), limiting the maximum observation to 48 months. These filters left us with 835,766 monthly bills. Finally we only kept households with at least 6 months of monthly billing data. After applying this last filter, we obtained 766,308 monthly billing data or 72% of all monthly bills provided. We consider each of our conditions to impact households randomly; that is we believe that our filters do not systematically influence the households that remain within our sample. For example, households that could not be merged with census data are equally likely to be from higher or lower income groups.

[4]It takes at least 20 kWh to keep a refrigerator turned on for a month. The average monthly consumption for our sample is 560 kWh, and monthly consumption that is 10 times larger than the average is potentially an erroneous record.

3.4.2 Variables

Our key dependent variable is monthly kilowatt-hour (kWh). Table 3.1 shows that the average monthly use is 559.7 kWh. The key explanatory variables are the three most commonly used payment methods. Out of 766,308 monthly bills, 304,347 (or 40%) of those were paid using mailed checks. This is the most popular option across all the different payment methods. The second most common option is through electronic funds transfer, which accounts for 21% of all bills paid. Electronic funds transfer is often done through a third-party provider such as a bank. The next most popular payment type is followed by online payments at 11%. We lump electronic funds transfer and online payments into one category because they are both done electronically, but are not automated. The next most popular payment is automatic payment services at 10%. The remainder of the bills were paid through pay stations (6.8%), phone (3.6%), and cash (3.4%). PG&E refers to automated payment services as set it and forget it (PG&E 2008). Users may pay using one or more of the different payment types.

PG&E provided the smart meters for free, as an opt-out program. The data contains information on when PG&E sent a letter to individual households announcing access to smart meters, and we use this to create a binary variable that takes one if the bill is received after this campaign and zero otherwise. We confirmed with PG&E that the decision when to install the smart meter is determined by the utility company, not the household. The activation of smart meters is therefore an exogenous variable; this means that the household did not determine when the smart meter service begins. We exploit the timing of when smart meters were activated to examine whether the interaction of having access to real-time feedback and different payment methods is associated with a decrease in energy use. Although it is an opt-out program, we discuss possible issues related to this in Section 8. Some households in the sample never received the smart meter by the end of the study, and some had the smart meter at the beginning of the study. The average number of months billed with a smart meter is 13.4

In addition we include the following control variables: weather[5], late payments, and participation in the California Alternate Rates for Energy (CARE) program. Although it is unlikely that users pay the bill in a way that depends on the weather, we control for weather variation.

We use information on the bill date and the payment date to determine late payments. A payment is late if the households pays after the due date (21 days after the bill date). We do not consider payments made after 31 days of the due date as late, as those households are subject to termination, but those are infrequent events. Table 3.1 shows that 17.3% of all bills are paid late.

Low-income households can participate in the California Alternate Rates for Energy (CARE) program. There are eligibility requirements to receive lower energy rates. For example, households with 2 or less and who earn $32,040 or less are eligible for reduced rates. We observe periods when households enroll in or out of these programs. In our data, 22.1% of bills had reduce rates during our study period.

We also collected census block information on median income, percent of households with college degrees, median age, percent of homeowners, average share of household below 18, share of houses build after 2000, and the share of houses that uses electric heating. The sample of bills come from three different regions. Inland Hills account for 38% of all the bills, followed by the Coast at 36.2%; the remainder of the sample come from Central Valley.

[5]This consists of summer indicator variables, the interaction of summer indicator variables and the three geographic locations, and year dummy variables. We also have information in heating and cooling degree days from the nearest weather station for each census block, and using this to control for weather does not change the conclusion of our results.

Table 3.1: Summary statistics.

Variables that change over time in the sample	Mean	St. Dev.	Min	Max
Monthly kWh	559.7	442.3	20	4998
Bills paid by check (binary)	0.408	0.492	0	1
Bills paid online (binary)	0.329	0.47	0	1
Bills paid automatically (binary)	0.092	0.289	0	1
Bills paid by cash (binary)	0.035	0.183	0	1
Bills paid by pay station (binary)	0.068	0.251	0	1
Bills paid by phone (binary)	0.035	0.185	0	1
Months billed with smart meter* (binary)	13.48	11.44	0	48
Periods after smart meter announcement (binary)	0.12	0.325	0	1
Heating degree days (divided by 100)	2.33	1.839	0	13.13
Cooling degree days (divided by 100)	0.618	1.122	0	7.902
Late payments (binary)	0.173	0.378	0	1
Enrolled in CARE**	0.221	0.415	0	1
Variables that do not change over time				
Median income (census)	73,910	36,050	6,917	250,000
Median age (census)	38.6	8.852	13.9	81.3
Fraction with college degrees (census)	0.363	0.227	0	1
Fraction of homeowners (census)	0.595	0.261	0	1
Fraction of households below 18 (census)	0.223	0.075	0	0.466
Fraction of homes built after 2000 (census)	0.099	0.138	0	0.989
Fraction of homes with electric heating (census)	0.222	0.12	0	0.796
Age of the account with utility	13.15	12.08	0	55
Ever logged on the dashboard (binary)	0.362	0.481	0	1

Notes: These observations are at the bill level. This data is based on an unbalanced panel where the minimum number of monthly household observation is 6 months and the maximum number of monthly observation is 48. The average number of observation per household is 37. *This variable is only examined in robustness tests in the online appendix. **California Alternate Rates for Energy program for low-income households.

3.5 Methods

Our goal is to test whether different payment methods influence electricity use. We define $\ln(kWh_{it})$ to be the natural log of the total monthly electricity consumption in kilowatt-hours (kWh) for household i in period t. Let $mailing\ check_{it-1}$ be a binary variable that takes value 1 if household i paid the previous month's bill by mailing a check. We include household fixed effects because we want to focus on how the variation in energy use changes when they switch to a different payment method.

$$\ln(kWh_{it}) = \alpha_i + \beta_1 \times mailing\ check_{it-1} + \epsilon_{it}. \tag{3.1}$$

We test equation (3.1) using different payment methods by replacing the binary variable $mailing\ check_{it-1}$ with another payment type. Households may choose different payment methods for a variety of reasons such as convenience or avoiding late payments. It is unlikely that households change the way they pay their bill to increase or decrease their energy use. It is possible that household characteristics that influence payment type may also influence energy use. To address potential issues arising from this, we provide various robustness test in Section 7.

To test our second hypothesis, we estimate the interaction of the introduction of smart meters with different payment methods. The binary variable, $Smart\ meter_{it}$ equals one starting in periods when the utility company sent campaign letters to households to access their smart meters and zero otherwise. We estimate equation (3.2), where the vector $\boldsymbol{\gamma}'\boldsymbol{\theta}$ is a shorthand to capture weather and year controls.

$$\ln(kWh_{it}) = \alpha_i + \beta_1 Smart\ meter_{it-1} + \beta_2 Payment\ method_{it-1} +$$

$$\beta_3 Smart\ meter_{it-1} \times Payment\ method_{it-1} + \boldsymbol{\gamma}'\boldsymbol{\theta} + \epsilon_{it}. \tag{3.2}$$

There are possible issues of self-selection that may bias our estimates, so we provide additional tests to try and eliminate potential concerns related to this. We provide several robustness tests in Section 7 to address potential bias in our results.

3.6 Results

The results of regression equation (1), which test how type of bill payment influences energy use, are provided in Table 3.2. Model (1) shows that housholds use 1.6% ($p < 0.01$) less electricity on average in periods when they pay by check compared to other payment types. Model (2) shows that online payments are associated with 0.8% increase in energy use compared to other payment types, but this estimate is weakly significant ($p < 0.10$). Model (3) shows that users who enroll in automated payment services (APS) use 4.6% ($p < 0.05$) more energy compared to periods when they use other payment types. Model (4) shows the results when both dummy variables for online and automatic payments are included, and the estimates for those variables are similar to those in models (2) and (3). Although the economic impact of our estimate may be modest for an individual user, extrapolating the results to millions of households is sizable. We provide a discussion on the size and implications of this impact in Section 8.

Table 3.2: Fixed effects (within) regression results of payment method on monthly electricity consumption.

Dependent variable: Monthly $\ln(kWh)$ at time $t+1$				
How users pay at time t	(1)	(2)	(3)	(4)
Paid by check	-0.016***			
	(0.004)			
Paid online		0.008*		0.010**
		(0.004)		(0.004)
Paid automatically			0.046**	0.049**
			(0.021)	(0.021)
Household fixed effects	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Observations	745,543	745,543	745,543	745,543
R-squared	0.035	0.035	0.035	0.035
Number of households	20,765	20,765	20,765	20,765

Notes: Cluster-robust standard errors in parentheses; $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$. Weather controls include an indicator variable for summer, year and the interaction of those two with dummy variables for the Coast, the Inland Hills, and the Central Valley. An alternative weather control using heating and cooling degree days gives similar results.

3.6.1 Payment types and smart meters

Table 3.3 provides the results of regression equation (2), which tests our second hypothesis on the interaction between electronic payments and access to daily cost and consumption information on electricity usage. We focus on the interaction of payment type and smart meter installation. Households who pay by check (not electronic; not automated) reduce their energy use by 1.5% on average after the utility company announced the installation of the smart meter. In contrast, we find that households who pay online tend to increase their energy use by 1.1% on average after access to smart meters were publicized, but this is weakly statistically significant ($p < 0.10$). This shows that access to smart meters is not always associated with a decrease in energy use, and it may even have the opposite effect. One possible reason for this is that if the convenience provided by online payments may make users less likely to review their bills more closely. In contrast, we find that those who pay using checks decrease their energy use even further by 1.5% ($p < 0.01$) in periods when they have access to smart meters.

Households who enroll in automated ways to pay reduce their energy use by 4.2% ($p < 0.01$), on average, in periods after the campaign on smart meters. This suggests that having access to more granular information may make energy cost and consumption more salient, leading to an attenuation of the increase in energy use associated with automated payments.

Table 3.3: Fixed effects (within) regression results of smart meters and different payment methods on monthly electricity consumption.

Dependent variable: Monthly $\ln(kWh)$ at time $t+1$			
	(1)	(2)	(3)
Paid by check	-0.015^{***} (0.004)		
Paid by check $\times$ Smart meters	-0.015^{***} (0.006)		
Paid online		0.007 (0.004)	
Paid online $\times$ Smart meters		0.011^{*} (0.006)	
Paid automatically			0.050^{**} (0.021)
Paid automatically $\times$ Smart meters			-0.042^{***} (0.011)
Smart meters	-0.000 (0.004)	-0.010^{**} (0.004)	-0.002 (0.004)
Household fixed effects	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Observations	745,543	745,543	745,543
R-squared	0.035	0.035	0.035
Number of households	20,765	20,765	20,765

Notes: Cluster-robust standard errors in parentheses; $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

Across the three models in Table 3.3, only model (2) shows a decrease in average energy use of 1% in periods after households received a mail about their smart meters. We implemented the Dubin-Wu-Hausman test for endogeneity, and we do find evidence that suggests that online payments are endogenous and is associated with the introduction of smart meters. This suggests that the treatment of providing information may not be as effective or even has an opposite impact when users are more focused on the convenience of paying online as opposed to having access to more granular information. This means that providing access to information may not be sufficient when users prioritize the convenience of paying online, making information costs less salient despite providing them more frequent access to this information.

3.7 Robustness Tests

The type of payment method users prefer may be correlated with other household characteristics that also influence energy use. For example, if higher income households are more likely to switch to electronic or automated payments then some of the changes in energy use may be attributed to income as opposed to the type of payment. We address potential sources of bias in three ways. First, we look at the interaction of payment type with the CARE program to examine if the results are consistent or if it varies by income. Second, we estimate adoption models of different payment types using a logit model, then we use propensity score matching as a robustness test whether the payment method can explain part of the variation in energy consumption. Third, we repeat the test using the random effects model that allows us to include time-invariant variables such as income and education. If the results remain consistent after adding various census-level controls then that would provide more evidence that the type of payment method is associated with changes in energy use and less likely to be driven by these household and dwelling characteristics.

3.7.1 Does the impact differ by income?

The impact of smart meters may be different for low- versus high-income households using information on CARE enrollment. Those who have CARE receive reduced rates during that period. Table 3.4 shows the results with the inclusion of enrollment to the CARE program and its interaction with smart meters. The three models in Table 3.4 show that average energy use increases by about 1-1.5% ($p < 0.01$) in periods when households received reduced rates. The interaction of CARE and smart meter is associated with about a 2% increase in energy use ($p < 0.01$). This means that the impact of smart meters is not always associated with a decrease in energy use. We are not the first to observe that providing additional information may have the opposite effect to energy conservation (see Asensio and Delmas 2015). We discuss this in more detail in Section 8.

The interaction of CARE and the three different payment types are not significantly different than zero. This suggest that the impact of salience from online payments and automation are not likely to apply to lower income households because they may continue to closely follow their bill despite the added convenience of these services. This suggests that the impact of payment methods and real-time feedback may vary by income.

Table 3.4: The association of payment types and the introduction of smart meters and its interaction with enrollment into CARE for reduced rates.

Dependent variable: Monthly ln(kWh) at time $t+1$			
	(1)	(2)	(3)
CARE	0.015**	0.010*	0.012**
	(0.006)	(0.006)	(0.006)
CARE × Smart meter	0.020***	0.023***	0.019***
	(0.007)	(0.007)	(0.007)
Smart meter	-0.006	-0.018***	-0.008**
	(0.004)	(0.005)	(0.004)
Paid by check	-0.012***		
	(0.004)		
Paid by check × CARE	-0.011		
	(0.007)		
Paid by check × Smart meter	-0.014**		
	(0.006)		
Paid online		0.005	
		(0.005)	
Paid online × CARE		0.005	
		(0.008)	
Paid online X Smart meter		0.014**	
		(0.006)	
Paid automatically			0.051**
			(0.022)
Paid automatically × CARE			-0.010
			(0.018)
Paid automatically × Smart meter			-0.039***
			(0.011)
Household fixed effects	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Observations	745,543	745,543	745,543
R-squared	0.035	0.035	0.035
Number of households	20,765	20,765	20,765

Notes: Cluster-robust standard errors in parentheses; $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

3.7.2 Logit models of selection

We obtain propensity scores from logit models to construct a covariate-balanced sample (see Becker et al. 2002 for more details on propensity-score models). The estimates from equation (3.3) are in Table 3.5:

$$\textit{ever paid online}_i = \alpha + \boldsymbol{\kappa}'\boldsymbol{\Omega} + \boldsymbol{\nu}'\boldsymbol{\phi} + \epsilon_i, \tag{3.3}$$

where $\boldsymbol{\Omega}$ is a vector of household characteristics that includes census demographics information, geography, visiting a website, fraction of times late, and $\boldsymbol{\phi}$ is a vector of census building characteristics such as the fraction of homes with one or less bedrooms, the fraction of homes with electric heating, and the fraction of homes built after 2000.

Model (1) shows the results of the mode logit model where the dependent variable equals one if a household ever paid online and zero otherwise. Model (2) shows the results for when the dependent variable equals one if a household ever paid automatically and zero if not. The controls include whether the household ever visited their personalized dashboard[6], the fraction of months they were late in paying the bill, the fraction of months they received reduced rates through the CARE program, and several census block household and dwelling characteristics.

The results of the logit models are reported as marginal effects, and several household and dwelling characteristics are associated with online and automated ways to pay. Visits to the dashboard, the fraction of households with a college degrees in a census block, the average household size, and the fraction of dwellings with less than two beds are positively associated with online payments ($p < 0.01$). The fraction of months with late payments, the fraction of months with re-

[6]This personalized webpage contains daily cost and consumption information in periods after they have access to smart meters. Users can visit their personalized dashboard even before the introduction of smart meters, but this site does not contain daily cost and consumption information.

duced rates, the age of the account, and the fraction of homeowners in a census block are negatively associated with the adoption of online payments ($p < 0.05$). The age of the account, the fraction of college degrees in a census block, the fraction of homes with electric heating, and geography are positively associated with enrollment in automatic payments ($p < 0.10$). Visits to a personal dashboard, the fraction of months with late payments, the fraction of months with reduced rates, the average household size, and the fraction of dwellings with two beds or less are negatively associated with enrollment in automatic payments ($p < 0.10$).

Table 3.5: Marginal impact of demographics on payment type of logit models used in propensity score matching.

Dependent variable:	Ever paid online	Ever paid automatically
	(1)	(2)
Ever visited a website	0.444^{***}	-0.006^{*}
	(0.007)	(0.004)
Fraction of months with late payment	-0.092^{**}	-0.050^{*}
	(0.046)	(0.029)
Fraction of months in CARE	-0.071^{***}	-0.017^{***}
	(0.011)	(0.006)
Income quartile 2 (binary)	0.005	−0.003
	(0.012)	(0.006)
Income quartile 3 (binary)	0.029^{**}	−0.002
	(0.015)	(0.007)
Income quartile 4 (binary)	0.031	0.003
	(0.019)	(0.008)
Age of account	-0.011^{***}	0.002^{***}
	(0.000)	(0.000)
Fraction of census block with college degrees	0.126^{***}	0.097^{***}
	(0.032)	(0.013)
Fraction of census block that are homeowners	-0.057^{**}	0.007
	(0.025)	(0.011)

Notes: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. Regression results continued in the next page.

Continuation of regression results from Table 3.5.

Dependent variable:	Ever paid online	Ever paid automatically
	(1)	(2)
Average fraction oh household below 18	−0.110	−0.077
	(0.101)	(0.047)
Fraction of dwellings with less than two beds	0.126***	−0.090***
	(0.042)	(0.020)
Fraction of dwelling with electric heating	0.002	0.062***
	(0.044)	(0.020)
Coast (binary)	0.013	0.012*
	(0.014)	(0.007)
Inland Hills (binary)	0.016	0.014**
	(0.012)	(0.006)
Fraction of dwelling built after 2000	0.020	−0.013
	(0.031)	(0.016)
Observations	20,765	20,765
Log Likelihood	−11,637.920	−5,832.716
Note:	*p<0.1; **p<0.05; ***p<0.01	

3.7.3 Propensity score matching

We use the results of logit models and use nearest neighbor as a measure of identifying controls (see Becker et al. 2002 for more discussion on nearest neighbor). Figure 3.1 shows the difference in the standardized bias across the covariates between the unmatched and matched samples for those who pay online. We find that the standardized bias in the unmatched sample are mostly between −25% and 25% with the exception of ever visiting the dashboard and the age of the account. The standardized bias did decrease for these two variables for the matched sample. Figure 3.2 shows the difference in the standardized bias between the unmatched and matched samples for those who enroll in automated payment services. We find that standardized bias before the matching is greater compared to the matched sample. The range of the standardized bias before matching is between −30% to 50%, and this is reduced to the range between −10% and 10% with the matched sample.

Figure 3.1: Plot of standardized percent bias across covariates before and after matching for those who paid online.

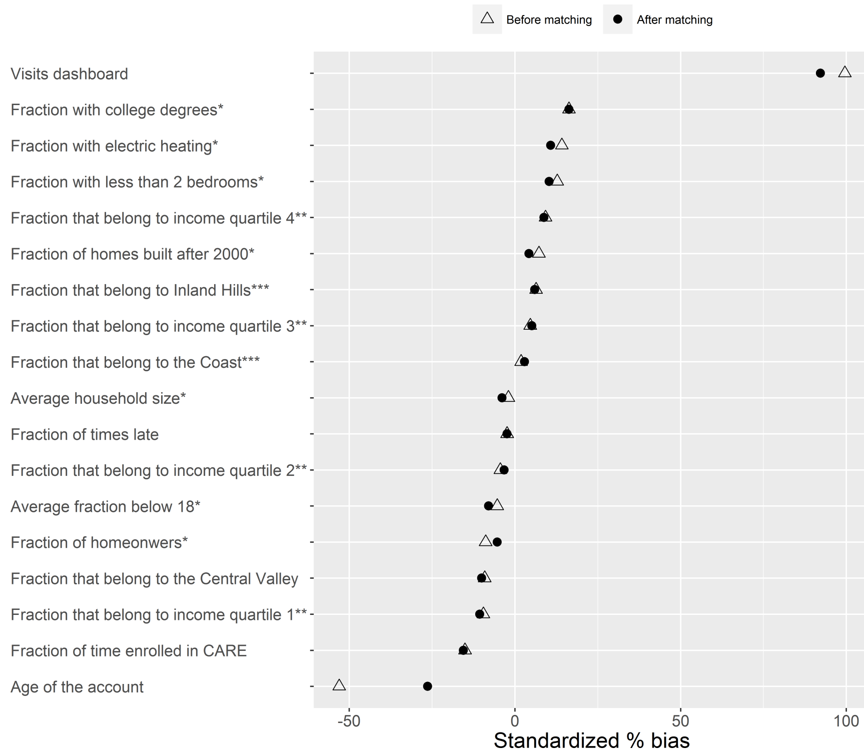


Figure 3.2: Plot of standardized percent bias across covariates before and after matching for those who paid automatically.

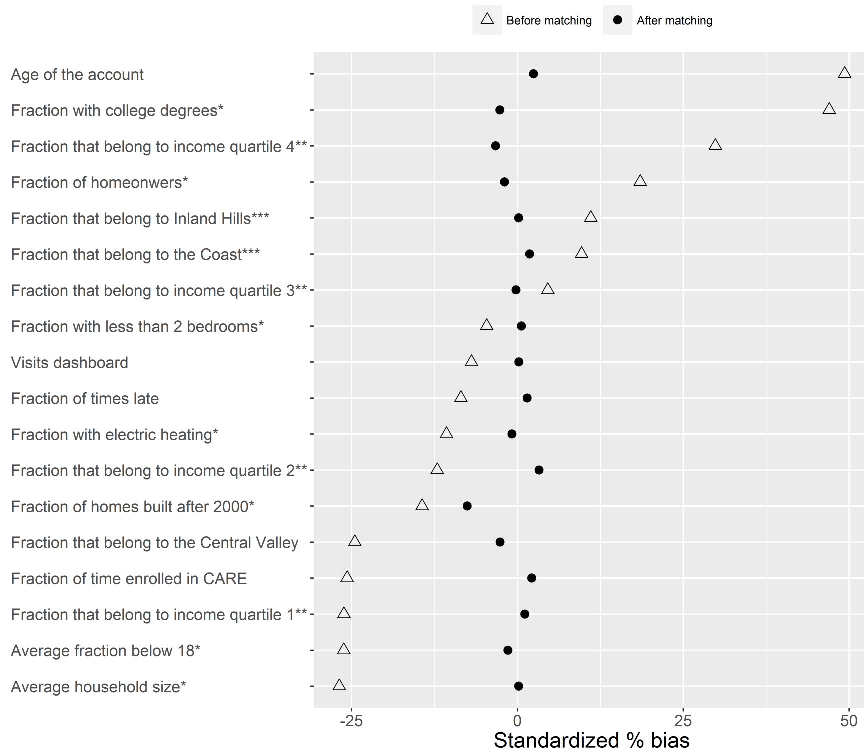


Table 3.6 summarizes the results for the propensity-matched sample for those who pay online or automatically. Model (1) in Table 3.6 shows that users who pay online are associated with a 0.6% average increase in energy use, but this is not statistically significant at the 0.10 level. Model (2) shows that users who pay using automated payment services is associated with a 5.0% increase in average electricity use. This supports our hypothesis that switching to automated ways to pay is associated with an increase in energy use. We do not include a model where users who pay by check are matched because we consider this as the pool of users that are matched with those who pay online or automatically. The interaction of paying automatically and publicizing smart meters is associated with a 2.5% ($p < 0.05$) decrease in average energy use. This supports our hypothesis that access to real-time feedback can mitigate the increase in energy use associated with automatic payments.

Table 3.6: Results of propensity-matched sample of households who pay online and automatically.

Dependent variable: Monthly $\ln(kWh)$ *at time* $t+1$	Online payments (1)	Automatic payments (2)
Paid online	0.006	
	(0.004)	
Paid online × Smart meters	0.008	
	(0.005)	
Paid automatically		0.050**
		(0.021)
Paid automatically × Smart meters		-0.025**
		(0.011)
Smart meters	-0.013***	-0.006
	(0.004)	(0.009)
Household fixed effects	Yes	Yes
Weather controls	Yes	Yes
Observations	682,080	146,769
R-squared	0.038	0.033
Number of households	19,324	3,686

Notes: Cluster-robust standard errors in parentheses; $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

3.7.4 Random effects models

The results of the random effects with several census block household and dwelling controls are consistent with results from the fixed effects models. We find that users who pay online use 1.1% more energy on average in periods after smart meters were publicized, but this is weakly statistically significant ($p < 0.10$). Users who enroll in automated payments reduce their energy use by 4.0% ($p < 0.01$) in periods after the smart meters were announced. The results in Model (3) support Hypothesis 2 that the interaction of automated ways to pay the bill and access to smart meters is associated with a decrease in energy use.

Table 3.7: Results of random-effects regression of monthly electricity use on different payment types, household characteristics and the introduction of smart meters.

	Dependent variable: Monthly ln(kWh) at time $t+1$		
	(1)	(2)	(3)
Paid by check	-0.015^{***}		
	(0.004)		
Paid by check × Smart meter	-0.014^{**}		
	(0.006)		
Paid online		0.005	
		(0.004)	
Paid online × Smart meter		0.011^{*}	
		(0.006)	
Paid automatically			0.042^{**}
			(0.019)
Paid automatically × Smart meters			-0.040^{***}
			(0.012)
Received mail about smart meters	0.000	-0.010^{**}	-0.002
	(0.004)	(0.004)	(0.004)
Late payments	Yes	Yes	Yes
Ever visited a website	Yes	Yes	Yes
Enrollment in CARE	Yes	Yes	Yes
Income brackets	Yes	Yes	Yes
Education brackets	Yes	Yes	Yes
Homeowner brackets	Yes	Yes	Yes
Below 18 brackets	Yes	Yes	Yes
Age brackets	Yes	Yes	Yes
Built after 2000	Yes	Yes	Yes
Electric heating	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Observations	745,543	745,543	745,543
Number of households	20,765	20,765	20,765

3.8 Discussion, Limitations, and Conclusions

3.8.1 Different payment types are associated with changes in energy use

The purpose of this paper is to examine whether different payment types may make energy cost more or less salient. We compared changes in energy use when users pay online by visiting a website or when they enroll in automated payment services. These two payment types make it more convenient for users to pay their energy bill, but one is automated and the other is not. We find that users who switch to paying online or through automated payments increase their energy consumption by 0.8% and 4.6%. In contrast, we find evidence that users who pay their bill by mailing checks is associated with a 1.6% decrease in energy use compared to periods when they use alternative ways to pay their bill. The results show that different payments methods may influence changes in energy use. This may mean that electronic ways to pay the bill may make energy costs or consumption less salient leading to an increase in energy use.

Users may process cost and consumption differently based on how their pay their utility bills. One possible explanation is that the process of writing the check and mailing it to the utility company may be associated with higher attention and better recall compared to paying online or automatically. In contrast, paying bills by going online or using automated payment services may make energy bills less salient because households may take less time to review the bill or not even bother to look at them at all. Electronic and automated ways to pay the bill are designed to make payments faster and hassle free, but they may have unintended consequences for conservation behavior. We do not suggest that these services should be abolished, but there are potential strategies that may mitigate this impact.

Providing real-time feedback may make energy cost and consumption more salient, leading to a reduction in energy use, but this impact may vary by demographics and how users pay their bill. We find that those who pay by check reduce an additional 2% of their energy use in periods when they have access to smart meters, but this is not the case for those who pay online. Information about cost and consumption do not always lead to a reduction in energy use, and in some cases it may even have the opposite effect. For instance, Asensio and Delmas (2015) find that households that were given financial information were more likely to increase energy use compared to households who were given information about the impact of electricity on our health and the environment.

Here we find that the same type of information may have a different impact depending on how users pay their bill. Our results suggest that access to real-time feedback may not be sufficient if the convenience of paying online does not encourage users to pay more attention to their energy use.

Those who pay automatically are likely to do so because they are worried about missing a payment, but this may be associated with less attention to the bill. We find that access to real-time feedback is associated with a decrease in energy use for those who pay automatically. This suggest that the causal impact of real-time feedback may be more heterogeneous than what earlier studies have shown. This means that providing access to real-time feedback should be complemented with various programs that could be tailored-fit to certain demographics and how they choose to interact with the utility company.

Users who receive reduced rates respond to real-time feedback differently. We find that those who enroll in the CARE program increase their energy use during those periods because of reduced rates. We expected that having access to smart meters can provide them with greater savings by helping them find opportunities to reduce their energy use, but we found the opposite. We find that those with reduced rates were associated with a further increase in energy use in periods with access to real-time feedback. This means that additional information does not always translate into decreased energy use. One possible explanation for this is that users may consider electricity prices to be low enough to curb their conservation efforts because it is not likely to create substantial savings.

The size of the impact of switching from paper to automated payments may look small for individual users, but this can be substantial in aggregate. To put the environmental implications in perspective, switching from paper to automated bills is roughly a 24kWh (or 4.4% of 560kWh, the average monthly electricity use in our sample) increase in monthly energy use. This may look small, but 24 kWh multiplied by 540,000 households (assuming 10% of the population switches to automated payments) is about 13 million kWh per month, which is equivalent to 1.5 million metric tons of carbon emissions per year in California. This is equivalent to generating roughly 30% of monthly operating capacity (65 MW) of a single coal power plant in California, the Argus Cogeneration plant (the only coal power plant in California at the time this paper was written). A 100% adoption of automated payments is likely to lead to a significantly higher impact.

There are significant benefits (mostly through cost reduction) in digitizing service operations, but this increase in convenience may make consumption cost less salient for some users. For example, some households may be less likely to read their energy bills after enrolling into automated payment services. Although this increase in consumption is modest in size (∼4.4%), the collective impact for more than one million households is significant. However, we find evidence to support that access to daily information may influence some users to decrease this energy cost. This implies that efforts to increase convenience may lead to a slight increase in consumption, but utility companies who are interested in managing demand can mitigate this unintended consequence by providing their consumers with equally convenient ways to access their own historical and real-time data.

3.8.2 Limitations and Future Directions

We discuss several limitations to our study, and provide directions for future research. In this study, we used field data instead of experimental data, and thus may have potential issues of endogeneity. Although there are many household factors that can influence how users pay that we cannot control for, our study focuses on how users change their payment types over time. In addition, our robustness test using different estimators, the random effects and propensity-matched samples, are consistent with our earlier findings. Yet, even after controlling for several household factors the only way to provide evidence for causality is to conduct a randomized controlled experiment. One possible direction is to conduct lab experiments that will test whether different ways to pay may influence the salience of cost or consumption. For example, users who use convenient, quick ways to pay their bill may recall fewer details about the bill compared to those who pay using alternative methods. This will provide more evidence that different payment methods may influence the salience of energy costs.

Households can opt-out of not having a smart meter and continue to pay additional monthly fees to continue services on an analog meter. We do not have information on which households opted out in our program, but we do have households in the sample that never get a smart meter at the end of the study. This could be because the utility has not yet installed smart meters at those locations or that the household opted out. If users who think they will not benefit from the smart meters opt out, then our results should show a stronger impact of smart meters on energy use because those who choose to receive the smart meter may be indifferent or they believe they will benefit from it. Since our results show that the impact of smart meters can be positive depending on demographics and how users pay their bill suggest that this bias is not as strong if at all present. Future studies can examine why some users are averse to the adoption of new technologies such as smart meters and prefer to pay additional fees to continue service using analog meters.

Appendices

Appendix A Appendix to Operational Response to Climate Change

A.1 CDP Survey Questions and Sample Firms

Table A.1 summarizes the list of CDP questions we use each year to construct the data used in this study. CDP switched the order of some of the questions asked, but they provided us with a reference guide to each question across the years included in this study. Copies of the survey questionnaire are available online (https://www.cdp.net/Documents/Guidance/2014/Climate-change-reporting-guidance-2014.pdf, p. 48, last accessed 10-29-2016). In the next section, we describe the filters and how we determine the sample.

Table A.1: Source of information in the CDP survey from 2010 to 2014.

	CDP survey year				
Information	2010	2011	2012	2013	2014
Investment Cost	**9.7** Please use table below to describe your company's actions to reduce its GHG emissions.	**3.3** Did you have emissions reduction initiatives that were active within the reporting year?	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**CC3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.
Annual Savings	**9.7** Please use table below to describe your company's actions to reduce its GHG emissions.	**3.3** Did you have emissions reduction initiatives that were active within the reporting year?	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**CC3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.
Emissions savings	**9.7** Please use table below to describe your company's actions to reduce its GHG emissions.	NA	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**CC3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.
Type	NA	**3.3** Did you have emissions reduction initiatives that were active within the reporting year?	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.	**CC3.3b** For those initiatives implemented in the reporting year, please provide details in the table below.
Driver	NA	**3.3b** What methods do you use to drive investment in emissions reduction activities?	**3.3c** What methods do you use to drive investment in emissions reduction activities?	**3.3c** What methods do you use to drive investment in emissions reduction activities?	**CC3.3c** What methods do you use to drive investment in emissions reduction activities?

The bold numbers represent the CDP survey question to obtain information used for that year. Information on lifetime is obtained using a combination of the data and the summary statistics from "Lower emissions, higher RIO: the rewards of low carbon investments" (CDP 2014).

Table A.2 shows a sample of global firms that report to CDP, their country, GICS Industry and the number of times they reported from 2010 to 2014. We handpicked 35 US firms to illustrate the diversity of industries that report to CDP and the frequency of how often they report. The 10 global firms are arbitrarily chosen to show the variation in the countries that report to CDP.

Table A.2: Sample of 35 handpicked US firms and 10 arbitrarily chosen global firms.

Company name	Country	GICS Industry	Times reported
Alcoa Inc.	USA	Metals & Mining	5
Amgen, Inc.	USA	Biotechnology	4
Apache Corporation	USA	Oil, Gas & Consumable Fuels	5
AT&T Inc.	USA	Wireless Telecommunication Services	5
Best Buy Co., Inc.	USA	Specialty Retail	4
Boeing Company	USA	Aerospace & Defense	5
Cisco Systems, Inc.	USA	Communications Equipment	5
Clorox Company	USA	Household Products	5
Colgate Palmolive Company	USA	Personal Products	5
Coca-Cola Company	USA	Beverages	5
Dow Chemical Company	USA	Chemicals	5
FedEx Corporation	USA	Air Freight & Logistics	5
Gap Inc.	USA	Specialty Retail	5
General Electric Company	USA	Industrial Conglomerates	5
Google Inc.	USA	Internet Software & Services	5
Goodyear Tire & Rubber Company	USA	Auto Components	4
The Hertz Corporation	USA	Road & Rail	4
Intel Corporation	USA	Computers & Peripherals	5
Levi Strauss & Co.	USA	Textiles, Apparel & Luxury Goods	4
NVIDIA Corporation	USA	Semiconductors & Semiconductor Equipment	5
Office Depot, Inc.	USA	Specialty Retail	5
Oracle Corporation	USA	Software	3
Pall Corporation	USA	Machinery	5
PG&E Corporation	USA	Multi-Utilities	5
Philip Morris International	USA	Tobacco	5

Company name	Country	GICS Industry	Times reported
Southwest Airlines Co.	USA	Airlines	5
Starbucks Corporation	USA	Hotels, Restaurants & Leisure	5
Target Corporation	USA	Multiline Retail	5
Walt Disney Company	USA	Media	5
Wal-Mart Stores, Inc.	USA	Food & Staples Retailing	5
Weyerhaeuser Company	USA	Real Estate Investment Trusts (REITs)	5
Xerox Corporation	USA	Office Electronics	5
H&R Block Inc	USA	Diversified Consumer Services	3
Kimball Office	USA	N/A	2
Xcel Energy Inc.	USA	Multi-Utilities	5
Krug Inc.	Canada	Household Durables	3
Chaun-Choung Technology Corp	Taiwan	Computers & Peripherals	2
Goldcorp Inc	Canada	Metals & Mining	5
Korean Air	South Korea	Airlines	4
SABMiller	United Kingdom	Beverages	5
BillerudKorsns	Sweden	Paper & Forest Products	2
Intermediate Capital Group	United Kingdom	Capital Markets	3
Mahindra & Mahindra Financial Services	India	Consumer Finance	2
Royal Bank of Canada	Canada	Commercial Banks	3
BP	United Kingdom	Oil, Gas & Consumable Fuels	5

A.2 The Sample and Metrics of Attractiveness

We discuss how we determine the sample for our main analysis on payback periods and MACs. Our analysis on payback periods uses these project-level data from 2010 to 2014. We divide the investment costs by the annual savings for each project reported on the survey questions listed on Table A.1. We describe the filters and the number of projects excluded at each stage. There are 14,621 projects with investment cost and savings information. (This includes projects from firms that report only once.) We remove any project that has a payback period greater than 100 as we believe those to be erroneous (e.g. adding additional zeros or misplacing decimal places for costs). This left us with 14,314 observations. We also remove projects with a positive investment cost and payback periods below 0.01. For instance, we would remove a project with a $100 cost but payback period of 0.0001 because we believe those to be possible typos (e.g. adding additional zeros or misplacing decimal places for savings), but we keep projects with zero cost and payback period of zero because many firms explicitly write zero for the cost. Adding this filter left us with 14,107 observations. Examples of projects with no investment costs include reducing waste and energy use through behavioral means or adjusting processes. Although these projects may include transaction costs that may not be reflected in what they report, firms implement these projects based on the values they mention. We remove all projects with a Cook's statistic greater than 0.0001 to avoid projects with high leverage that are potential outliers. Cook's statistic is obtained by regressing the payback period on the year. See Faraway (2014) for a discussion on Cook's statistic. (The typical cutoff used is $2p/n$ where p is the number of parameters and n is the total number of observations.) This left us with 13,249 observations. Because we want to estimate firm-level time trends, we require that firms report at least twice during this period. This resulted in an unbalanced sample of 978 firms and 11,941 projects for our analysis on payback periods. We applied the same filters for determining the sample used for marginal abatement costs with a couple of differences. We remove the top and bottom 5% of MACs (above $226 and below $-399 per ton of CO2e) from the sample. We consider this, just like payback periods greater than 100 years, to be erroneous or possible outliers. We exclude projects that have a Cook's statistic greater than 0.0002, slightly higher than for payback periods because our sample is smaller. This resulted in an unbalanced sample of 8,165 projects across 808 firms.

We explain how we calculate MACs for implemented projects. The necessary information to calculate MACs is available beginning in 2010 except in 2011 because CDP did not ask carbon emissions reduction for individual projects in that year. Firms report annual carbon emissions reduction, investment costs I, and annual monetary savings S. We convert all costs and savings to USD using annual exchange rates published by the International Monetary Fund (http://data.imf.org, last accessed January 16, 2015). We use the lifetime of each project L based on the average lifetime by type as reported in CDP. We also assume an interest rate of $r = 20\%$ such that $\delta = 1/(1 + r)$. Interest rates of 20% and above are commonly found in the energy efficiency literature (Hausman 1979; Hassett and Metcalf 1993), so we use similar estimates in calculating MACs. The marginal abatement cost of a single opportunity is calculated as follows (Laitner et al. 2003, p. 6-110):

$$MAC = \frac{I - \sum_{t=0}^{L} \delta^t S}{emissions\ reduction^* L}. \tag{A.1}$$

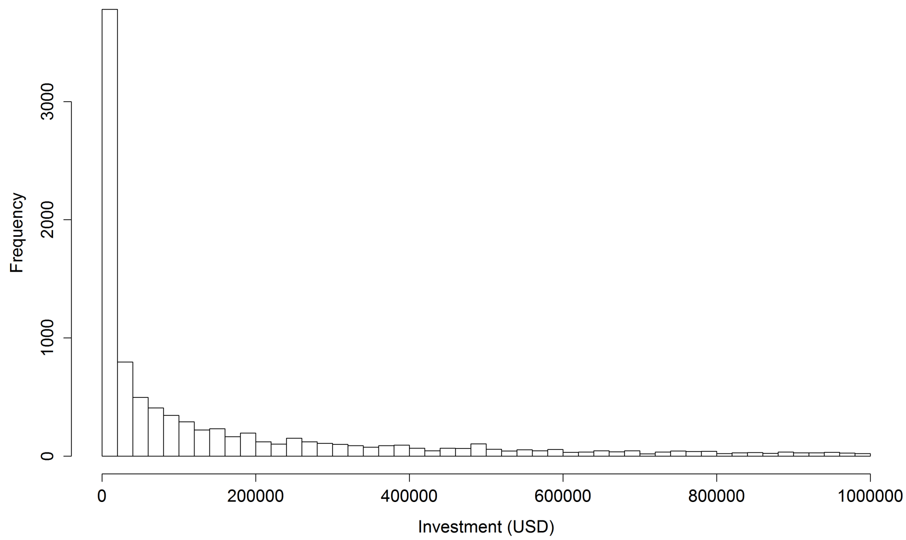
We comment on the distribution of the size, measured using the investment cost (in USD), of projects implemented. The histogram in Figure A.1 shows the distribution of investments costs (in USD) under \$1 million (N=9,219). There are 6,106 projects that have an investment cost greater than \$100,000, and 8,767 projects have an investment cost greater than \$10,000. If we consider projects greater than \$100,000 to be large, then the histogram shows that firms pursue both small and large projects.

A.3 The Variation in the Trends

We estimated the average trends in payback and MACs using a mixed-effects model. Let the metric of interest be denoted by y_{kit} for project k, firm i, at time t:

$$y_{kit} = \beta_i \times year_t + \alpha_i + \eta_{kit}, \tag{A.2}$$

Figure A.1: Histogram of investment costs (in USD) under $1 million.



where β_i is the individual trend of interest, $year_t$ is the linear trend, α_i are the random intercepts, and η_{kit} are the residuals. We use the mixed-effects model to predict individual time trends $\hat{\beta}_i$. The best linear unbiased predictor of individual time trend is given by

$$\hat{\beta}_\mathbf{i} = \beta_{GLS} + DZ_i'(Z_i D Z_i' + R)^{-1}(y_i - X_t \beta_{GLS}), \tag{A.3}$$

where X_t is the data of the independent variables, D is the variance of the individual trends, Z_i is data on individual firms, and R is the variance of the error term η_{kit}. The intuition behind equation (A.3) is the individual time trend β_i is equal to the average trend β_{GLS} plus a term that captures how the individual firm departs from that average.

A.4 Data Related to Different Types of Opportunities

In this section, we describe the different types of projects firms can pursue to reduce their carbon emissions. Table A.3 shows the average payback period by types by GICS Sector. (Recall that each of the 66 GICS Industries is classified into one of the 10 GICS Sectors.) The project type is chosen by the survey taker from a drop-down menu provided by CDP. The description of the carbon abatement types is on page 49 of the CDP survey guide (CDP 2014). The descriptions of these activities were the same in all the years in this study. We provide descriptions and examples here.

We begin by describing the first four types with the shortest payback period. (1) “Behavioral change” are projects associated with energy use awareness and training employees to minimize waste. These projects have the lowest payback period of 1.01 year, and they account for 5.2% of the projects (617 out of 11,941). (2) “Transportation use” and (3) “transportation fleet” are activities that reduce emissions associated with the movement of goods and services. These include route optimization, switching old vehicles with more energy efficient ones, changing the mode of transportation (e.g. air to sea; truck to rail). Transportation use and fleet have an average payback period of 1.30 and 1.96. Transportation use and fleet are 2.4% and 4% of all projects reported. (4) “Product design” are projects related to changes in packaging or innovations that enable downstream users to avoid carbon emissions. The average payback period for those types is two years, and they account for 1.2% of all projects.

We describe the next four types with the shortest payback period. (5) “Process emissions reduction” are initiatives related to manufacturing, e.g., the purchase of new equipment, changes in operations, process materials selection, etc. (CDP 2014, p. 49). The average payback period for these projects is 2.02 years, and they account for 5.3% of the projects. (6) “Energy efficiency: processes” are projects that reduce the required input fuel related to the production of goods and services. This type includes emissions reduction through changes in the chemical or physical processes other than fuel combustion such as emissions from calcination in cement manufacturing, emissions from aluminum smelting, or emissions from catalytic cracking in petrochemical processing. These projects have an average payback of 2.32 years and account for 27.8% of reported projects. Energy efficiency: processes is the second most frequently reported type. (7) "Low carbon purchase" involve procuring energy from low-carbon emitting generators. The average payback period for this type is 2.92, and they account for 1.1% of all projects. (8) "Fugitive emissions reduction" are ways to reduce emissions related to capturing agricultural methane release or preventing oil, natural gas or refrigerant leaks. The average payback for this type is 3.11, and they account for 0.06% of projects. This is the least reported type.

The three remaining types have the longest payback period. (9) "Energy efficiency: building services" are opportunities related to the more efficient use of buildings such as installing building controls and HVAC (heating, ventilation, and air conditioners). The average payback period of these types is 3.18, and this is the most frequently reported type (29.8%). "Energy efficiency: building fabric" are opportunities related to building insulation. The average payback period of this type is 3.96, and this type accounts for 5.7% of the projects. "Low carbon install" are opportunities that are related to installing low-carbon energy generators (e.g. wind and solar powered generators). This type has the longest average payback period at 5.53 years, and this is 6.4% of the projects firms report.

Table A.3: Average payback period by Global Industry Classification Standard (GICS) and type.

GICS Sector	Behavioral change	EE: Building fabric	EE: Building services	EE: Processes	Fugitive emissions reductions	Low carbon energy installation	Low carbon energy purchase	Process emissions reductions	Product design	Transportation: fleet	Transportation: use
Consumer Discretionary	0.78	4.44	2.73	2.07	4.10	4.22	2.58	2.00	4.52	1.54	1.56
Consumer Staples	0.57	3.55	3.18	2.67	3.44	5.58	2.63	1.42	1.19	2.46	0.74
Energy	1.34	4.07	2.77	2.91	3.55	6.71	0.10	2.61	3.34	1.92	2.90
Financials	0.73	4.65	3.90	2.19	2.09	6.20	0.08	2.29	1.12	1.21	1.99
Health Care	1.75	2.75	3.36	2.78	3.90	6.86	2.64	2.05	2.97	0.42	0.28
Industrials	1.43	4.37	3.22	2.22	2.72	5.78	4.28	1.47	1.91	2.41	0.91
I.T.	0.59	3.03	2.32	1.73	1.54	4.35	3.57	1.48	0.94	1.72	1.08
Materials	0.58	3.31	2.69	2.12	2.94	4.85	2.98	2.73	2.24	1.35	0.30
Telecom Services	0.53	5.33	3.08	3.17	1.60	4.31	1.02	2.98	3.94	1.29	0.88
Utilities	3.10	5.60	4.55	2.23	2.76	6.41	0.31	0.81	NA	3.18	1.91
Overall average	1.01	3.96	3.18	2.32	3.11	5.53	2.92	2.02	2.00	1.96	1.30
Total observations	617	679	3,555	3,324	83	763	127	634	140	474	289

Notes: CDP did not ask firms to report the type in 2010. There are 859 projects reported in 2010. The remaining 36 was simply left blank by the survey taker. We do not use 2010 in the analysis for fraction of opportunities by type, but the trends capture the payback periods during this period. We assume that firms are likely to select and report roughly the same fraction of opportunities that are directly related to core and those that are not.

A.5 Classification and refinements of core-aligned Opportunities

There are opportunities that are directly related to the core production processes of a company and some that are not, but this classification is not unambiguous. In this section, we describe how we classified projects as being core-aligned or not. We ran our analysis with several iterations of this classification and found that our conclusions are robust to these various refinements. We classify all 11 opportunity types for each of the 66 GICS Industries, a more granular classification than GICS Sectors, on whether they are directly related to core company operations or not. We (the three authors) individually give a score from one to five, where one represents projects that we confidently identified as *not* directly related to core operations and five as opportunities that we confidently believe to be directly related to the core operations in that industry. We consider that we have a disagreement in our classification if the maximum difference between the scores from any two of us is greater than one. After our first round of classification, we identified where we disagreed. We discussed why we coded differently and independently reclassified these pairs. A project is directly related to core operations if it receives an average score of 4 or higher across the three authors.

We provide summary statistics of our disagreement rates between the first and second round in Table A.4. We weight the agreement by the number of projects reported for each GICS Industry and carbon abatement type. We discuss the rate of agreement for the first round. The rate of agreement is highest for low carbon energy purchase and low carbon energy installation at 96% and 88%. Product design and transportation fleet and use are the three types with the lowest rate of agreement at 53%, 64%, and 72%. The (weighted) average agreement of the first round is 75%. We discussed and reclassified opportunities where we disagreed. The overall agreement rates increased by 10%, but there are still many projects where we could not come to a consensus. There was a large increase in agreement for process emissions reduction and EE: processes from 77% to 96% and from 79% to 91%. One of the authors refined all remaining disagreements and reclassified 1,442 projects as either core-aligned or not. Although we scored all behavioral projects as three, one of the authors reclassified them using firm-level information. We show all the results at different stages of the classification later in this document.

Table A.4: Weighted agreement rate on 66 × 11 GICS Industry and type matrix classification.

Type	Percent (weighted) agreement First round	Percent (weighted) agreement Second round
Transportation: fleet	64%	72%
Transportation: use	72%	79%
Low carbon energy purchase	96%	96%
Process emissions reductions	77%	96%
Fugitive emissions reductions	78%	93%
Product design	53%	65%
Energy efficiency: Processes	79%	91%
Energy efficiency: Building services	71%	81%
Energy efficiency: Building fabric	75%	82%
Low carbon energy installation	88%	92%
Percent (weighted) agreement:	75%	85%

We report the inter-rater reliability, the fraction of opportunities that two raters give the same score. In the first round, author A and author B agreed on 77.9%, author A and C agreed on 74.8%, and author B and C agreed on 72.2% across their scores. In the second round, author A and author B agreed on 80.9%, author A and C agreed on 75.8% and author B and C agreed on 81.0% across their scores.

Hypothesis 3 predicts that firms that pursue a higher fraction of opportunities that are directly related to core operations will experience more favorable payback trends relative to the average trend. We summarize the regression results across different refinements in this section. We present the association of the fraction of projects directly related to core operations (or core-aligned projects) on payback and MAC trends at each stage of our coding process using different sample filters. The direction of the results is consistent in each round of coding, using different filters.

Table A.5 summarizes the association of the fraction of core projects on payback period trends. The dependent variable is the firm-level time trends $\hat{\beta}_i$. Models (1) and (2) summarize the results for the first round of coding each industry and activity type pair. Model (1) excludes all projects that are ambiguous. This sample includes 3,761 core-aligned projects and 4,602 that are not. Model (2) excludes all observations where we had a disagreement. This model includes 3,401 core-aligned projects and 4,513 that are not. Models (3) and (4) summarize the results for the second round of refinement. Model (3) excludes ambiguous projects and classification where we had disagreements.

This model includes 4,395 core-aligned projects and 4,685 projects that are not. Model (4) is a refinement of Model (3) where all disagreements were classified by one of the authors based on the description of the projects. This model includes 4,563 core-aligned projects and 5,224 that are not. The estimates of the association of core-aligned projects across the four models vary, but they are all negative and statistically significant at the 0.10 level.

Table A.5: The association of core-aligned opportunities on payback trends.

	Dependent variable: payback time trend			
	weighted OLS			
	(1)	(2)	(3)	(4)
Fraction of opportunities directly related to core operations	-0.066^{***}	-0.053^{***}	-0.037^{*}	-0.037^{**}
	(0.020)	(0.019)	(0.019)	(0.019)
GICS Sector	Yes	Yes	Yes	Yes
Observations	908	869	944	962
R^2	0.085	0.081	0.084	0.073

Notes: *p<0.1; **p<0.05; ***p<0.01. heteroscedastic-robust standard errors are in parentheses. All models exclude behavioral projects because they were considered ambiguous. The results in the main paper include behavioral projects, and the conclusions remain the same.

We present the results on the association of core-aligned opportunities and MAC trends. We repeat the analysis in the previous section and replacing the dependent variable with the trends in marginal abatement costs. Models (1) − (4) in Table A.6 use the same filters in the same order as the models in Table A.5. We find no evidence that MAC trends are changing over time, but we find significant heterogeneity in the trends. Table A.6 shows no evidence that pursuing a higher fraction of core-aligned opportunities is associated with deteriorating (increasing) MAC trends.

Table A.6: The association of core-aligned opportunities on MAC trends.

	Dependent variable: MAC trends			
	(1)	(2)	(3)	(4)
Fraction of opportunities directly related to core operations	−0.437	−0.374	−0.150	−0.089
	(0.300)	(0.263)	(0.232)	(0.228)
GICS Sector	Yes	Yes	Yes	Yes
Observations	742	695	778	793
R^2	0.019	0.024	0.023	0.023

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroscedastic-robust standard errors are in parentheses.

A.6 Additional Material for Robustness Tests

A.6.1 Robustness tests on the trends using fixed and random effects estimators

We conduct our analysis on the trends using alternative estimators with the fixed and random effects models. We construct a panel dataset by taking the mean of payback and MACs for each firm each year so that we have one observation per year per firm. The results of estimating the time trends are summarized below. Table A.7 shows that the estimates for the time trend of payback periods for the fixed (Model (1)) and random (Model (2)) effects are very similar to the estimates of the mixed effects model at 0.078 and 0.085. We find no evidence that MACs are changing over time for both fixed and random effects (Models (3) and (4)). We reach the same conclusions with the results using the mixed effects models.

Table A.7: Fixed and random effects models of the payback period and MAC trends.

Dependent variable:	*mean payback*$_{it}$		*mean MAC*$_{it}$	
	(1) Fixed effects	(2) Random effects	(3) Fixed effects	(4) Random effects
Time trend	0.078^{**}	0.085^{**}	0.253	0.180
	(0.038)	(0.036)	(0.605)	(0.578)
Constant	–	2.595^{***}	–	-29.408^{***}
	–	(0.145)	–	(2.434)
Observations	3,167	3,167	2,274	2,274

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.6.2 Robustness tests on firms with and without a dedicated budget for energy efficiency

We explore whether our result is consistent with firms with and without a dedicated budget for implementing energy efficiency, but some do not. There are 705 firms that mention they have a dedicated budget at any point during the study. Model (1) in Table A.8 shows that firms with a budget are associated with a 0.095 increase in average payback period (or a 3.4% increase per year; $p-$value<0.01), but Model (2) shows no evidence that payback periods are increasing for firms that do not have a budget. Models (3) and (4) in Table A.8 show no evidence that MACs are increasing or decreasing for firms with and without a budget.

Table A.8: Mixed-effects results of payback period and MAC time trends for firms that do and no not mention a budget for carbon abatement.

	Dependent variable:			
	Payback period		Marginal abatement costs	
	(1)	(2)	(3)	(4)
	Firms with a budget	Firms without a budget	Firms with a budget	Firms without a budget
Time trend	0.095^{***}	0.002	0.438	-0.371
	(0.035)	(0.056)	(0.571)	(0.958)
Constant	2.569^{***}	2.640^{***}	-30.177^{***}	-28.402^{***}
	(0.138)	(0.226)	(2.372)	(3.825)
Number of firms	705	273	590	218
Total projects	8980	2961	6244	1921

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.6.3 Robustness tests on firms that belong to high-emissions sectors

The availability of core-aligned projects may vary among sectors that emit a larger amount of Scope 1 and 2 (direct and indirect carbon emissions), so we repeat the analysis with firms that belong to sectors that account for a significant portion of total carbon emissions. It is possible that firms that belong to Financials and I.T. may have fewer opportunities to reduce emissions that are directly related to their core operations. On the other hand, firms that belong to high emissions industries such as Materials and Industrials may have more such opportunities. Based on the 2014 CDP report, the six sectors with the largest total Scope 1 and 2 emissions are Utilities, Industrials, Materials, Consumer Staples, Consumer Discretionary and Energy (CDP 2014, p. 10). Table A.9 shows that the results are consistent with earlier ones, but the magnitude is larger. The models chosen in Table A.9 use the same filters as those in Table A.5, i.e., model (1) in Table A.9 is the first round of coding, excluding projects that are ambiguous. We confirm our earlier findings that firms that pursue a higher fraction of core-aligned opportunities experience more favorable trends compared to those that do not.

Table A.9: The association of core-aligned projects and payback period for energy-intensive sectors.

	Dependent variable: payback time trend			
	(1)	(2)	(3)	(4)
Fraction of opportunities directly related to core operations	-0.077^{***}	-0.061^{**}	-0.080^{***}	-0.073^{***}
	(0.026)	(0.024)	(0.025)	(0.023)
GICS Sector	Yes	Yes	Yes	Yes
Observations	565	556	583	599
R^2	0.081	0.073	0.082	0.069

Notes: *p<0.1; **p<0.05; ***p<0.01. Heteroscedastic-robust standard errors are in parentheses.

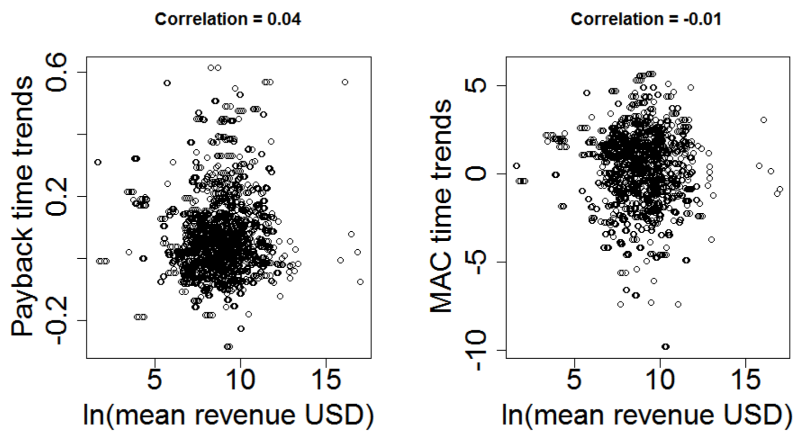
A.6.4 Robustness tests with alternative ways to identify potential outliers

We repeat our analysis on the trends of payback periods and MACs using Tukey's range test within each firm. This approach considers values that are 1.5 times below or above the interquartile range IQR (the third quartile q_3 minus the first q_1) as possible outliers; in other words values below $q_1 - 1.5 \times IQR$ or above $q_3 + 1.5 \times IQR$ are potential outliers. We implement Tukey's range test as an alternative to using Cook's statistic to identify potential outliers. We also test models where we implement both tests to identify potential outliers. In all cases, we find no statistical significance ($p > 0.10$) in the estimates of payback period and MAC trends.

A.6.5 Robustness tests related to firm size and project scale

We find weak evidence that the size of the firm is associated with payback and MAC trends. Figure A.2 shows plots of payback time trends, MAC time trends, and firm size as measured by mean revenue in USD. We take the natural log of mean revenue to reduce skew and improve the fit. The correlation of payback time trends and the natural log of mean revenue is 0.04. The correlation of MAC time trends and the natural log of mean revenue is -0.01. We find little evidence that the size impacts the trends.

Figure A.2: Correlation plot of firm size and trends.



The trends of cost are decreasing over time, but the magnitude of this varies depending on the filter we implement. Table A.10 shows the trends using mixed effects models where the dependent variable is the cost of a project divided by the firm's average cost across all projects from 2010 to 2014. In model (1) we do not implement any filters, but we could not achieve convergence in our estimates using Stata. This could be due to the wide variation and potential outliers in cost even after normalization. A measure of a potential outlier in one dimension (e.g. cost) may not necessarily be an outlier in another after a transformation (e.g., payback period) (Fraway 2005 p. 67). In model (2), we remove the top 10% of all observations as we consider those potential outliers. The results in model (2) show that the costs are decreasing by 6.3 percentage points. One possible limitation of removing the top 10% of observations is that it does not account for within firm variation. With a strict cutoff, a set of large projects for one firm may all be considered an outlier when in reality these projects are indeed legitimate. Model (3) shows the results after removing potential outliers identified using Tukey's range test for each firm. We find that the normalized cost is decreasing by −2.4 percentage points relative per year. (The dependent variable is in decimal form.) This decrease could be driven by a variety of factors such as learning or finding more smaller projects.

Table A.10: Trends in the cost of abatement projects after normalizing by each firm's average cost.

	Dependent variable: $cost_{kit}/\overline{cost_i}$		
	(1)	(2)	(3)
Trend	–	-0.063^{***}	-0.024^{**}
	–	(0.015)	(0.010)
Constant	–	1.234^{***}	1.096^{***}
	–	(0.059)	(0.040)
Number of observations	11,931	10,695	9,737

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is in decimal form. Model (2) removes the top 10% of all observations before normalizing. This does not take into account within firm variation. Model (3) implements Tukey's range test for potential outliers with firm. There are projects that do not have a currency exchange and are therefore dropped, but they are included in the payback analysis because that does not require information on exchange rates.

Table A.11 shows the trends in the number of projects reported. The total number of opportunities firms report increased by 0.27 per year (or roughly 9.6% more per year).

Table A.11: Trends in the number of projects.

Dependent variable:	Number of projects (OLS)
Time trend	0.271^{***}
	(0.056)
Constant	2.826^{***}
	(0.199)
Total firms	978
Total observations	3,167

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.6.6 Robustness tests using alternative measures of time: cumulative projects and cumulative emissions reduction

We examine cumulative learning as an alternative measure of time. We measure this using the cumulative number of projects implemented and the cumulative emissions reduction to capture a firm's past experience in carbon abatement. We take the average of payback period across all projects for firm i at time t, and we do the same for MACs. We estimate a fixed-effects model on the mean payback period and mean MACs on the natural log of the cumulative number of projects per year. We then estimate a fixed-effects model on the average MAC on the natural log of cumulative emissions reduction. Model (1) (and (2)) of Table A.12 shows no evidence that the total cumulative number of projects is associated with changes in the average payback period. We repeat this for MAC in models (3) and (4); we find no evidence that the cumulative number of projects is associated with mean MACs. We find evidence in model (5) that a 10% increase in the total cumulative emissions reduction is associated with a \$0.171 increase in average MAC ($p < 0.01$). It is expected that as total cumulative emissions increase then marginal abatement costs become less attractive over time, but the magnitude of this impact suggests that opportunities do not deteriorate quickly.

Table A.12: Fixed effects models of payback period and MAC on alternative measures of time.

Dependent variable:	mean payback$_{it}$		mean MAC$_{it}$		mean MAC$_{it}$
	(1)	(2)	(3)	(4)	(5)
Cumulative number of projects	0.011	–	0.028	–	–
	(0.008)	–	(0.140)	–	–
Ln(cumulative number of projects)	–	0.072	–	0.472	–
	–	(0.067)	–	(1.055)	–
Ln(cumulative emissions reduction)	–	–	–	–	1.717**
	–	–	–	–	(0.674)
Observations	3,167	3,167	2,274	2,274	2,274
Note:					*p<0.1; **p<0.05; ***p<0.01

A.7 Example of opportunities classified as core-aligned and those that are not

Table A.13 provides examples from the data of core-aligned projects and those that are not. We provide the company name, the GICS Industry, the type (as chosen from the drop-down menu provided by CDP), the classification of core-aligned, and the description provided by the company.

Table A.13: Examples from the data of core-aligned projects and those that are not.

	Company name	GICS Industry	Type	Core-aligned	Description
1	Royal Bafokeng Platinum Ltd	Metals & Mining	EE: Processes	1	Vanes installed on first phase main ventilation fans. This is for scope 2 emissions, It is a voluntary initiative. Expected lifetime of 20 years
2	Oceana	Food Products	EE: Processes	1	Boiler Efficiency Optimisation Project - Boiler 1 at Oceana Brands. Oceana Brands has scheduled the same optimisation project for each of its four boilers over the next 4 years. Costs shown here are for one boiler. scope 1, voluntary, 10 year life time
3	Alcoa Inc.	Metals & Mining	Process emissions reductions	1	Alcoa continues to aggressively implement efforts to reduce smelting process interruptions that result in the release of perfluorocarbon emissions. In 2011, 12 smelters further reduced their emissions by adopting best practices and/or installing better raw material (point) feeders.
4	The Westpac Group	Commercial Banks	EE: Building services	0	Sealed all leakage points and penetrations
5	Ercros	Chemicals	EE: Building services	0	Tarragona site: Cooling system improvement
6	PerkinElmer, Inc.	Life Sciences Tools & Services	EE: Building services	0	Chiller operation efficiency improvement. Optimized and standardized how multiple chillers are used to provide required cooling most efficiently.
7	Devon Energy Corporation	Oil, Gas & Consumable Fuels	EE: Processes	1	In 2009, Devon voluntarily initiated an Energy Efficiency and GHG Reduction Program in Devon's nine field districts based on the Canadian Association of Petroleum Producers (CAPP) Fuel Gas Best Management Practices (BMPs). This program is focused primarily on Devon's non-oil sands operations and is the primary mechanism currently in place to reduce GHG emissions and operating costs from these assets. In 2012, 34 projects were implemented as part of this initiative. Projects result in Scope 1 and 2 emission reductions. Anticipated duration of savings ranges from 5 to 10+ years depending on the project.

	Company name	GICS Industry	Type	Core-aligned	Description
8	British American Tobacco	Tobacco	Low carbon energy installation	0	Voluntary - Life Span 5 - 10 Years: Replacement of the 72 high bay 500W sodium lights within the PMD area in Building 18 with energy efficient LED lighting running at 140W (Scope 1)
9	Cummins Inc.	Machinery	EE: Building services	0	Replace boiler, transformers and lighting and eliminate coal-fired furnace — Chonqing Cummins Engine Company, Chonqing, China – Engine Business Unit Note that CCEC is a Cummins joint venture company; we do not include JV emissions in our GHG goals, but we do include them in all associated activities.
10	Blue Coat Systems, Inc.	Electronic Equipment, Instruments & Components	Product design	1	Reduction in size of product packaging, removal of printed colateral from shipping boxes
11	Arcelor Mittal South Africa Ltd	Metals & Mining	EE: Processes	1	The Vanderbijlpark plant 30MW turbine was upgraded in 2011 and generate 6 to 10 MW additional to the average of the previous year (22 MW) depending on off gas availability. This means that an additional 52560 to 87600 tons of CO2 pa will be saved.
12	Merck & Co., Inc.	Pharmaceuticals	EE: Building services	0	Improvements to boiler systems at several sites around the world including; Variable Frequency Drive (VFD) replacements, stack energy recovery, economizers and improved insulation.
13	Santos Ltd	Oil, Gas & Consumable Fuels	EE: Processes	1	Energy efficiency measures at Fairview: Reduce number of operating compressors in Process Control System (PCS) (Scope 1, voluntary initiative).
14	Derwent London	Real Estate Investment Trusts (REITs)		0	Implement the findings of the detailed energy surveys at the three pilot buildings: Oliver's Yard, Johnson Building and Davidson Building

	Company name	GICS Industry	Type	Core-aligned	Description
15	Eni	Oil, Gas & Consumable Fuels	EE: Processes	1	Investments and operational initiatives in the refinery of Taranto. The value of the savings was estimated assuming 500 '/toe (fuel oil) and 10 /tCO2; a tax charge of 40% was considered. The savings are scope 1 and were evaluated on annual basis considering standard operation. All energy efficiency initiatives are voluntary, even if some of them were included in a programatic agreement with Italian Ministry of Environment they were anyway not mandatory. The lifetime of the initiatives is variable, ranging from a few years to until the plants will be in operation.
16	J Sainsbury Plc	Food & Staples Retailing	EE: Building services	0	Energy Efficiency Programme: Depot, 2 depots (Scope 1 & 2, Voluntary) Undertake detailed surveys and major energy efficiency works for lighting, heating and refrigeration into our distribution depots.
17	Crown Ltd	Hotels, Restaurants & Leisure	EE: Building services	1	Crown Melbourne - Building control optimisation will provide significant energy savings with the existing heating, cooling and ventilation systems.
18	Industrial Bank of Korea	Commercial Banks	EE: Building services	0	Introduction of PEAK electrivity control system in IBK's main office and IT data center Scope type: 2 Voluntary action
19	Anglo Platinum	Metals & Mining	EE: Processes	1	IGV Project (Peak Clipping): The IGVs were installed with the purpose of saving electrical energy during peak times of the day. A total of 6.8MW was identified for clipping at the Anglo Platinum Amandelbult (2.1MW) and Anglo Platinum Rustenburg (4.7MW) sections. This is a voluntary initiative which results in Scope 2 emission reductions. The expected project life is 5 years.
20	First Solar Inc	Electrical Equipment	EE: Processes	1	Optimized the compressed-air consumption of the air knives at our manufacturing facility in Perrysburg by installing solenoid valves. Scope 2; voluntary.

Appendix B Appendix to Managing Safety-Related Disruptions

We provide results on the association of safety-related disruptions on electricity production, measured by the capacity factor. We then describe how we construct the instrumental variable and the results of that regression.

B.1 The association of safety-related events and electricity production

We measure the impact of safety-related disruptions on capacity factor, the total electricity produced divided by the maximum electricity that can be generated within the same period[1]. We merge monthly capacity factor data from Davis and Wolfram (2012b) with our dataset. Table B.1 shows the results of fixed effects models where the dependent variable is the monthly capacity factor and the key independent variable is the number of safety-related events for that month. Model (1) shows that one event is associated with a 0.074 ($p < 0.01$) decrease in capacity factor for that month. This association is robust to the inclusion of time trends in model (2) and when weighted by plant capacity in model (3).

[1] A capacity factor of 100% means that the power plant produced the maximum electricity it can generate over that time period.

Table B.1: What is the association between the number of monthly events and capacity factors?

	Dependent variable: Monthly capacity factor		
	(1)	(2)	(3)
Number of safety-related events	-0.074^{***}	-0.072^{***}	-0.071^{***}
	(0.002)	(0.002)	(0.002)
Reactor fixed effects	Yes	Yes	Yes
Time trend	Yes	Yes	Yes
Weighted by plant capacity	–	–	Yes
Observations	16,066	16,066	16,066
R^2	0.207	0.208	0.208

Notes: Cluster-robust standard errors in parentheses; *p<0.1; **p<0.05; ***p<0.01

Next, we examine whether the type of event may impact capacity factors differently. Table B.2 shows the results where the number of safety-related events is broken down into five different types. These events are the same ones we discuss in the main manuscript, and their description is the same here. We find that events associated with the degradation of equipment have the highest impact on capacity factor at -0.09 ($p < 0.01$), followed by system actuation at -0.08. These two types of events are often associated with a shutdown of the plant and thus have the highest impact on production. We find that events associated with a technical specification or inoperable equipment have a -0.05 and -0.06 impact on capacity factor for that month ($p < 0.01$). Events that may have prevented a safety equipment from performing has the lowest impact on capacity factor at -0.03, but is still significant at the 0.01 level. Although the impact may vary by type, the results show that safety-related events are negatively associated with electricity production.

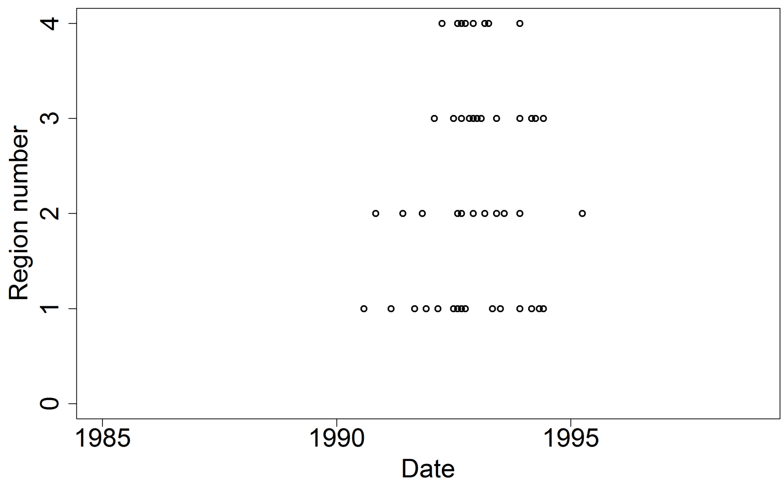
Table B.2: The association of different types of events on capacity factors.

	Dependent variable: monthly capacity factor		
	(1)	(2)	(3)
System actuation	-0.080^{***}	-0.074^{***}	-0.071^{***}
	(0.003)	(0.003)	(0.003)
Technical specification	-0.053^{***}	-0.052^{***}	-0.052^{***}
	(0.003)	(0.003)	(0.003)
Degradation	-0.094^{***}	-0.099^{***}	-0.102^{***}
	(0.006)	(0.006)	(0.006)
Inoperable	-0.055^{***}	-0.051^{***}	-0.050^{***}
	(0.014)	(0.014)	(0.014)
Prevent safety equipment	-0.033^{***}	-0.031^{***}	-0.027^{***}
	(0.007)	(0.007)	(0.007)
Reactor fixed effects	Yes	Yes	Yes
Time trend	Yes	Yes	Yes
Weighted by plant capacity	–	–	Yes
Observations	16,066	16,066	16,066
R^2	0.195	0.197	0.198
Note:	*p<0.1; **p<0.05; ***p<0.01		

B.2 Addressing potential endogeneity using an instrument

We constructed an instrument as the fraction of other reactors that have adopted PRA within a regulatory region. All nuclear plants belong to one of four Nuclear Regulatory Commission (NRC) region. Each point in Figure B.1 shows when a reactor adopted PRA and their NRC region. The plot shows that many reactors in regions 3 and 4 submitted within a short time window at the start of 1992. Most of those who submitted in 1990 and 1991 belong to regions 1 and 2. We find a strong correlation (0.92) on the fraction of other reactors within a regulatory region that have adopted with the focal reactor's adoption decision, but we have little reason to believe that the decision of other reactors to adopt PRA has a direct influence on the number of safety-related disruptions at another plant. These two conditions makes the variable we constructed a promising instrument.

Figure B.1: Adoption of individual reactors by regulatory region



Model (1) in Table B.3 shows that PRA adoption is associated with a -0.91 ($p < 0.01$) decrease on the average number of safety-related events when we use an instrument. This magnitude decreases to -0.50 ($p < 0.01$) in model (2) when we add a time trend. The estimate is -0.44 ($p < 0.01$) when we add weights using the capacity of the plant in model (3). Although the estimates vary, they are negative and statistically significant at the 0.01 level. The estimates here are also larger compared to our base estimates using Poission regression. Overall, this suggests that adoption of PRA is associated with a decrease in the number of monthly events even when we try to account for potential endogeneity issues using instruments.

Table B.3: Results of instrumental variables regression of PRA adoption on monthly events.

	Dependent variable: Monthly number of events		
	(1)	(2)	(3)
PRA adopt	-0.907^{***}	-0.498^{***}	-0.442^{***}
	(0.028)	(0.067)	(0.067)
Reactor fixed effects	Yes	Yes	Yes
Time trend	–	Yes	Yes
Weighted capacity	–	–	Yes
Observations	16,066	16,066	16,066
R^2	0.123	0.131	0.140
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Bibliography

Aflaki, S., Kleindorfer, P. R., Polvorinos, M., and Sáenz, V. (2013a). Finding and implementing energy efficiency projects in industrial facilities. *Production and Operations Management*, 22(3):503–517.

Aflaki, S., Kleindorfer, P. R., Polvorinos, M., and Sáenz, V. (2013b). Finding and implementing energy efficiency projects in industrial facilities. *Production and Operations Management*, 22(3):503–517.

Allcott, H. and Greenstone, M. (2012). Is there an energy efficiency gap? *Journal of Economic Perspectives*, 26(1):3–28.

Anderson, S. T. and Newell, R. G. (2004). Information programs for technology adoption: The case of energy-efficiency audits. *Resource and Energy Economics*, 26(1):27–50.

Angell, L. C. (2001). Comparing the environmental and quality initiatives of baldrige award winners. *Production and Operations Management*, 10(3):306–326.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Asensio, O. I. and Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences*, 112(6):E510–E515.

Azizi, M. M. (2014). PRA application to offshore drilling critical systems. *Probabilistic Safety Assessment and Management*, 12.

Baker, E., Clarke, L., and Shittu, E. (2008). Technical change and the marginal cost of abatement. *Energy Economics*, 30(6):2799–2816.

Beck, N. and Katz, J. N. (2007). Random coefficient models for time-series–cross-section data: Monte Carlo experiments. *Political Analysis*, 15(2):182–195.

Becker, S. O., Ichino, A., et al. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal*, 2(4):358–377.

Blanco, C., Caro, F., and Corbett, C. J. (2017). An inside perspective on carbon disclosure. *Business Horizons*.

Blumstein, C., Krieg, B., Schipper, L., and York, C. (1980). Overcoming social and institutional barriers to energy conservation. *Energy*, 5(4):355–371.

Boss, G., Gilbert, M., Olson, P., Reid, R., Cass, J., Fitzgerald, D., Swich, W., and Goff, J. (1997). NRCs oversight of nuclear power plants. Congressional Report.

Casella, G. and Berger, R. L. (2002). *Statistical inference*, volume 2. Duxbury Pacific Grove, CA.

CDP (2014a). 2014 CDP Report.

CDP (2014b). CDP climate change survey responses.

CDP (2014). CDP: Guidance for companies reporting on climate change on behalf of investors & supply chain members 2014. `https://www.cdp.net/Documents/Guidance/2014/Climate-change-reporting-guidance-2014.pdf`. Last accessed 10-19-2015.

Chatterji, A. K., Levine, D. I., and Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy*, 18(1):125–169.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *The American Economic Review*, 99(4):1145–1177.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd Edition)*. Routledge.

Cooke, R. M., Ross, H. L., and Stern, A. (2011). Precursor analysis for offshore oil and gas drilling: From prescriptive to risk-informed regulation. Technical report, Resources for the Future.

Corbett, C. J. and Kirsch, D. A. (2001). International diffusion of ISO 14000 certification. *Production and Operations Management*, 10(3):327–342.

Corbett, C. J. and Klassen, R. D. (2006). Extending the horizons: Environmental excellence as key to improving operations. *Manufacturing & Service Operations Management*, 8(1):5–22.

Corbett, C. J., Montes-Sancho, M. J., and Kirsch, D. A. (2005). The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Science*, 51(7):1046–1059.

Davis, L. W. and Wolfram, C. (2012a). Deregulation, consolidation, and efficiency: Evidence from us nuclear power. *American Economic Journal: Applied Economics*, 4(4):194–225.

Davis, L. W. and Wolfram, C. (2012b). Deregulation, consolidation, and efficiency: Evidence from US nuclear power: Dataset. *American Economic Journal: Applied Economics*, 4(4):194–225.

DeCanio, S. J. (1993). Barriers within firms to energy-efficient investments. *Energy Policy*, 21(9):906–914.

DeCanio, S. J. (1998). The efficiency paradox: Bureaucratic and organizational barriers to profitable energy-saving investments. *Energy Policy*, 26(5):441–454.

Delmas, M. and Toffel, M. W. (2004). Stakeholders and environmental management practices: An institutional framework. *Business Strategy and the Environment*, 13(4):209–222.

Delmas, M. A., Fischlein, M., and Asensio, O. I. (2013). Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy*, 61:729–739.

Faraway, J. J. (2014). *Linear models with R*. CRC Press.

Fine, C. H. (1988). A quality control model with learning effects. *Operations Research*, 36(3):437–444.

Fleiter, T., Schleich, J., and Ravivanpong, P. (2012). Adoption of energy-efficiency measures in SMEs-An empirical analysis based on energy audit data from Germany. *Energy Policy*, 51:863–875.

Freeman, M., Savva, N., and Scholtes, S. (2017). Economies of scale and scope in hospitals: Evidence of productivity spillovers across hospital services. *Working paper*.

Fremeth, A. R. and Holburn, G. L. (2012). Information asymmetries and regulatory decision costs: An analysis of US electric utility rate changes 1980–2000. *Journal of Law, Economics, and Organization*, 28(1):127–162.

Gerarden, T., Newell, R. G., and Stavins, R. N. (2015). Deconstructing the energy-efficiency gap: Conceptual frameworks and evidence. *American Economic Review*, 105(5):183–186.

Gilbert, B. and Zivin, J. G. (2014). Dynamic salience with intermittent billing: Evidence from smart electricity meters. *Journal of Economic Behavior & Organization*, 107:176–190.

Gillingham, K., Newell, R. G., and Palmer, K. (2009). Energy efficiency economics and policy. *Annual Review of Resource Economics*.

Goerlandt, F., Khakzad, N., and Reniers, G. (2016). Validity and validation of safety-related quantitative risk analysis: A review. *Safety Science*.

Gourville, J. T. and Soman, D. (1998). Payment depreciation: The behavioral effects of temporally separating payments from consumption. *Journal of Consumer Research*, 25(2):160–174.

Gray, J. V., Anand, G., and Roth, A. V. (2015). The influence of ISO 9000 certification on process compliance. *Production and Operations Management*, 24(3):369–382.

Greene, W. H. (2012). *Econometric Analysis (7th Edition)*. Boston, Pearson.

Gupta, S., Starr, M. K., Farahani, R. Z., and Matinrad, N. (2016). Disaster management from a POM perspective: Mapping a new domain. *Production and Operations Management*, 25(10):1611–1637.

Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3):1311–1353.

Harris, J., Anderson, J., and Shafron, W. (2000). Investment in energy efficiency: A survey of Australian firms.

Energy Policy, 28(12):867–876.

Hassett, K. A. and Metcalf, G. E. (1993). Energy conservation investment: Do consumers discount the future correctly? *Energy Policy*, 21(6):710–716.

Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *The Bell Journal of Economics*, pages 33–54.

Hendricks, K. B. and Singhal, V. R. (1997). Does implementing an effective TQM program actually improve operating performance? Empirical evidence from firms that have won quality awards. *Management Science*, 43(9):1258–1274.

Hendricks, K. B. and Singhal, V. R. (2005). An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. *Production and Operations Management*, 14(1):35–52.

Hickman, J. (1983). A guide to the performance of probabilistic risk assessments for nuclear power plants. *US NRC*.

Infotrends (2015). Will the check be in the mail? An examination of paper and electronic transactional mail.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *The American Economic Review*, 104(2):537–563.

Ittner, C. D. (1994). An examination of the indirect productivity gains from quality improvement. *Production and Operations Management*, 3(3):153–170.

Iyer, A., Saranga, H., and Seshadri, S. (2013). Effect of quality management systems and total quality management on productivity before and after: Empirical evidence from the Indian auto component industry. *Production and Operations Management*, 22(2):283–301.

Jackson, T. (1991). Least-cost greenhouse planning supply curves for global warming abatement. *Energy Policy*, 19(1):35–46.

Jacobs, F. R. and Chase, R. B. (2013). *Operations and supply chain management: The core*. McGraw-Hill.

Jaffe, A. B. and Stavins, R. N. (1994). The energy paradox and the diffusion of conservation technology. *Resource and Energy Economics*, 16(2):91–122.

Jain, N., Hasija, S., and Popescu, D. G. (2013). Optimal contracts for outsourcing of repair and restoration services. *Operations Research*, 61(6):1295–1311.

James, K. H. and Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, 1(1):32–42.

Jessoe, K. and Rapson, D. (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *The American Economic Review*, 104(4):1417–1438.

Jira, C. and Toffel, M. W. (2013). Engaging supply chains in climate change. *Manufacturing & Service Operations*

Management, 15(4):559–577.

Keller, W. and Modarres, M. (2005). A historical overview of probabilistic risk assessment development and its use in the nuclear power industry: A tribute to the late Professor Norman Carl Rasmussen. *Reliability Engineering & System Safety*, 89(3):271–285.

Kim, S.-H. and Tomlin, B. (2013). Guilt by association: Strategic failure prevention and recovery capacity investments. *Management Science*, 59(7):1631–1649.

King, A. and Lenox, M. (2002). Exploring the locus of profitable pollution reduction. *Management Science*, 48(2):289–299.

King, A. A. and Lenox, M. J. (2001). Lean and green? An empirical examination of the relationship between lean production and environmental performance. *Production and Operations Management*, 10(3):244–256.

Klassen, R. D. and McLaughlin, C. P. (1996). The impact of environmental management on firm performance. *Management Science*, 42(8):1199–1214.

Klassen, R. D. and Whybark, D. C. (1999). The impact of environmental technologies on manufacturing performance. *Academy of Management Journal*, 42(6):599–615.

Kleindorfer, P. R. and Saad, G. H. (2005). Managing disruption risks in supply chains. *Production and Operations Management*, 14(1):53–68.

Kleindorfer, P. R., Singhal, K., and Wassenhove, L. N. (2005). Sustainable operations management. *Production and Operations Management*, 14(4):482–492.

Kolk, A., Levy, D., and Pinkse, J. (2008). Corporate responses in an emerging climate regime: The institutionalization and commensuration of carbon disclosure. *European Accounting Review*, 17(4):719–745.

Kolk, A. and Pinkse, J. (2004). Market strategies for climate change. *European Management Journal*, 22(3):304–314.

LaChance, J. L., Martz, H., Anderson, D., Englehardt, M., Whitehead, D., and Wheeler, T. (2003). Handbook of parameter estimation for probabilistic risk assessment. Technical report, NUREG/CR-6823. Washington, DC: US NRC. ACC: MOL. 20060126.0121.

Laitner, J. A., Worrell, E., Galitsky, C., and Hanson, D. A. (2003). Characterizing emerging industrial technologies in energy models. *Lawrence Berkeley National Laboratory.*

Lapré, M. A., Mukherjee, A. S., and Van Wassenhove, L. N. (2000). Behind the learning curve: Linking learning activities to waste reduction. *Management Science*, 46(5):597–611.

Levine, D. I. and Toffel, M. W. (2010). Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. *Management Science*, 56(6):978–996.

Lochbaum, D. A. (2000). *Nuclear plant risk studies: Failing the grade.* Union of Concerned Scientists.

Mangen, A., Anda, L. G., Oxborough, G. H., and Brønnick, K. (2015). Handwriting versus keyboard writing: Effect on word recall. *Journal of Writing Research*, 7(2).

Markovitch, S. and Willmott, P. (2014). Accelerating the digitization of business processes. *McKinsey & Company, San Francisco.*

McKinsey & Company (2009). Pathways to a low-carbon economy. Last accessed 7-15-2015.

Meier, A. K. (1982). Supply curves of conserved energy. *Lawrence Berkeley National Laboratory.*

Muth, J. F. (1986). Search theory and the manufacturing progress function. *Management Science*, 32(8):948–962.

Muthulingam, S. and Agrawal, A. (2016). Does quality knowledge spillover at shared suppliers? An empirical investigation. *Manufacturing & Service Operations Management*, 18(4):525–544.

Muthulingam, S., Corbett, C. J., Benartzi, S., and Oppenheim, B. (2013). Energy efficiency in small and medium-sized manufacturing firms: Order effects and the adoption of process improvement recommendations. *Manufacturing & Service Operations Management*, 15(4):596–615.

National Commission on the BP Deepwater Horizon Oil Spill (2011). *Deep water: The Gulf oil disaster and the future of offshore drilling.* Perseus Distribution Digital.

Pachauri, R. K. and Allen, M. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Pacific Gas & Electric Company (2016). Pacific Gas and Electric Company Website Profile.

Pagell, M., Johnston, D., Veltri, A., Klassen, R., and Biehl, M. (2014). Is safe production an oxymoron? *Production and Operations Management*, 23(7):1161–1175.

Pil, F. K. and Rothenberg, S. (2003). Environmental performance as a driver of superior quality. *Production and Operations Management*, 12(3):404–415.

Poloski, J., Marksberry, D., and Atwood, U. C. (1999). Rates of initiating events at us nuclear power plants. Technical report, NUREG/CR-5750, US NRC, Washington (DC).

Prelec, D. and Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science*, 17(1):4–28.

Prelec, D. and Simester, D. (2001). Always leave home without it: A further investigation of the credit-card effect on willingness to pay. *Marketing Letters*, 12(1):5–12.

Rae, A., Alexander, R., and McDermid, J. (2014). Fixing the cracks in the crystal ball: A maturity model for quantitative risk assessment. *Reliability Engineering & System Safety*, 125:67–81.

Rajaram, K. and Corbett, C. J. (2002). Achieving environmental and productivity improvements through model-based process redesign. *Operations Research*, 50(5):751–763.

Rajaram, K., Jaikumar, R., Behlau, F., Van Esch, F., Heynen, C., Kaiser, R., Kuttner, A., and Van De Wege, I. (1999). Robust process control at Cerestar's refineries. *Interfaces*, 29(1):30–48.

Rees, J. V. (2009). *Hostages of each other: The transformation of nuclear safety since Three Mile Island*. University of Chicago Press.

Reid, E. M. and Toffel, M. W. (2009). Responding to public and private politics: Corporate disclosure of climate change strategies. *Strategic Management Journal*, 30(11):1157–1178.

Reiss, P. C. and White, M. W. (2005). Household electricity demand, revisited. *The Review of Economic Studies*, 72(3):853–883.

Repenning, N. P. and Sterman, J. D. (2001). Nobody ever gets credit for fixing problems that never happened: Creating and sustaining process improvement. *California Management Review*, 43(4):64–88.

Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science, 15–32*.

Rothenberg, S., Pil, F. K., and Maxwell, J. (2001). Lean, green, and the quest for superior environmental performance. *Production and Operations Management*, 10(3):228–243.

Rubin, E. S., Cooper, R. N., Frosch, R. A., Lee, T. H., Marland, G., Rosenfeld, A. H., and Stine, D. D. (1992). Realistic mitigation options for global warming. *Science*, 257(5067):148–149.

Science Based Targets (2016a). 200 Companies Commit to Science Based Targets, Surpassing Expectations for Corporate Climate Action. `http://sciencebasedtargets.org/2016/11/16/200-companies-commit-to-science-based-targets-surpassing-expectations-for-corporate-climate-action`. Last accessed 11-29-2016.

Science Based Targets (2016b). Science Based Targets Case Study: Coca-Cola Enterprises. `http://sciencebasedtargets.org/case-studies/case-study-coca-cola-enterprises`. Last accessed 11-29-2016.

Sexton, S. (2015). Automatic bill payment and salience effects: Evidence from electricity consumption. *Review of Economics and Statistics*, 97(2):229–241.

Shafir, E. and Thaler, R. H. (2006). Invest now, drink later, spend never: On the mental accounting of delayed consumption. *Journal of Economic Psychology*, 27(5):694–712.

Shin, J.-S. (1985). Perception of price when price information is costly: Evidence from residential electricity demand. *The Review of Economics and Statistics*, pages 591–598.

Smoker, T. J., Murphy, C. E., and Rockwell, A. K. (2009). Comparing memory for handwriting versus typing. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 53, pages 1744–1747. SAGE

Publications Sage CA: Los Angeles, CA.

Sodhi, M. S., Son, B.-G., and Tang, C. S. (2012). Researchers' perspectives on supply chain risk management. *Production and Operations Management*, 21(1):1–13.

Soman, D. (2003). The effect of payment transparency on consumption: Quasi-experiments from the field. *Marketing Letters*, 14(3):173–183.

Sønderhousen, U. (1993). Motivation for industry to increase energy efficiency. *Proceedings European Council for an Energy Efficient Economy*.

Sornette, D., Maillart, T., and Kröger, W. (2013). Exploring the limits of safety analysis in complex technological systems. *International Journal of Disaster Risk Reduction*, 6:59–66.

Speed, T. (1991). That BLUP is a good thing: The estimation of random effects: Comment. *Statistical Science, 42–44*.

Taylor, J. B., Wolak, F. A., et al. (2012). A comparison of government regulation of risk in the financial services and nuclear power industries. Technical report.

Thollander, P., Danestig, M., and Rohdin, P. (2007). Energy policies for increased industrial energy efficiency: Evaluation of a local energy programme for manufacturing SMEs. *Energy Policy*, 35(11):5774–5783.

US General Accounting Office (1985). *Probabilistic risk assessment: An emerging aid to nuclear power plant safety regulation*.

US NRC (1975). Reactor safety study: An assessment of accident risks in US commercial nuclear power plants. Technical report, United States Nuclear Regulatory Commission.

US NRC (1984). Licensee event report system, description of system and guidelines for reporting.

US NRC (1988). Individual plant examinations for severe accident vulnerabilities. *Generic Letter 88-20*.

US NRC (1991). Severe accident risks: An assessment for five US nuclear power plants. Technical report, Nuclear Regulatory Commission.

US NRC (1997). Individual plant examination program: Perspectives on reactor safety and plant performance. *NUREG-1560*.

Ward, D. (2014). The failure of marginal abatement cost curves in optimising a transition to a low carbon energy supply. *Energy Policy*, 73:820–822.

Worrell, E., Bernstein, L., Roy, J., Price, L., and Harnisch, J. (2009). Industrial energy efficiency and climate change mitigation. *Energy Efficiency*, 2(2):109–123.