

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Essays on Environmental Policy in Energy Markets

### Permalink

<https://escholarship.org/uc/item/05f4r67b>

### Author

Boomhower, Judson Paul

### Publication Date

2015

Peer reviewed|Thesis/dissertation

**Essays on Environmental Policy in Energy Markets**

by

Judson Paul Boomhower

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Severin Borenstein, Chair  
Associate Professor Lucas W. Davis  
Associate Professor Meredith Fowlie  
Professor Maximilian Auffhammer

Summer 2015

**Essays on Environmental Policy in Energy Markets**

Copyright 2015  
by  
Judson Paul Boomhower

## Abstract

Essays on Environmental Policy in Energy Markets

by

Judson Paul Boomhower

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Severin Borenstein, Chair

Producing and consuming energy involves costly environmental externalities, which are addressed through a wide range of public policy interventions. This dissertation examines three economic questions that are important to environmental regulation in energy. The first chapter measures the effect of bankruptcy protection on industry structure and environmental outcomes in oil and gas extraction. The second chapter measures additionality in an appliance replacement rebate program. Finally, the third chapter focuses on the environmental impacts of subsidizing electricity production from forest-derived biomass fuels.

The first chapter measures the incentive effect of limited liability. When liability is limited by bankruptcy, theory says that firms will take excessive environmental and public health risks. In the long run, this “judgment-proof problem” may increase the share of small producers, even when there are economies of scale. I use quasi-experimental variation in liability exposure to measure the effects of bankruptcy protection on industry structure and environmental outcomes in oil and gas extraction. Using firm-level data on the universe of Texas oil and gas producers, I examine the introduction of an insurance mandate that reduced firms’ ability to avoid liability through bankruptcy. The policy was introduced via a quasi-randomized rollout, which allows me to cleanly identify its effects on industry structure. The insurance requirement pushed about 6% of producers out of the market immediately. The exiting firms were primarily small and were more likely to have poor environmental records. Among firms that remained in business, the bond requirement reduced oil production among the smallest 80% of firms by about 4% on average, which is consistent with increased internalization of environmental costs. Production by the largest 20% of firms, which account for the majority of total production, was unaffected. Finally, environmental outcomes, including those related to groundwater contamination, also improved sharply. These results suggest that incomplete internalization of environmental and safety costs due to bankruptcy protection is an important determinant of industry structure and safety effort in hazardous industries, with significant welfare consequences.

The second chapter focuses on the importance of a regulator’s inability to distinguish between households responding to a subsidy, and households doing what they would also have

done in the absence of policy. Economists have long argued that many recipients of energy-efficiency subsidies may be “non-additional,” getting paid to do what they would have done anyway. Demonstrating this empirically has been difficult, however, because of endogeneity concerns and other challenges. In this paper we use a regression discontinuity analysis to examine participation in a large-scale residential energy-efficiency program. Comparing behavior just on either side of several eligibility thresholds, we find that program participation increases with larger subsidy amounts, but that most households would have participated even with much lower subsidy amounts. The large fraction of inframarginal participants means that the larger subsidy amounts are almost certainly not cost-effective. Moreover, the results imply that about half of all participants would have adopted the energy-efficient technology even with no subsidy whatsoever.

Finally, the third chapter addresses consequences of renewable energy subsidies in other markets. Electricity generated from logging residues provides a large and growing share of US renewable electricity generation. Much of the low-value wood used by biomass power plants might otherwise be left in the field. This increased harvest can negatively affect forest health. I investigate the supply of woody biomass fuel in Maine using a 15-year panel of prices and quantities for whole tree wood chips. I find that doubling the price of woody biomass increases harvest by about 64%. I also find that coal prices are a major determinant of woody biomass harvest. This suggests that environmental policies that raise the price of coal will affect forest health.

# Contents

<b>Contents</b>	<b>i</b>
<b>1 Bankruptcy, Insurance, and Environmental Risk</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Background . . . . .	3
1.3 Model . . . . .	6
1.4 Empirical Analysis and Data . . . . .	12
1.5 Results . . . . .	16
1.6 Discussion . . . . .	26
1.7 Conclusion . . . . .	28
<b>2 Additionality in Energy Efficiency Programs</b>	<b>49</b>
2.1 Introduction . . . . .	49
2.2 Conceptual Framework . . . . .	52
2.3 Program Background and Construction of Dataset . . . . .	55
2.4 Empirical Strategy . . . . .	57
2.5 Results . . . . .	59
2.6 Discussion . . . . .	63
2.7 Conclusion . . . . .	66
<b>3 Biomass Energy and Forests</b>	<b>76</b>
3.1 Introduction . . . . .	76
3.2 Model and instruments . . . . .	81
3.3 Data . . . . .	85
3.4 Results . . . . .	87
3.5 Conclusion . . . . .	89
<b>References</b>	<b>101</b>
<b>A Bankruptcy and Environmental Risk Appendix</b>	<b>110</b>
<b>B Additionality in Energy Efficiency Appendix</b>	<b>118</b>
<b>C Biomass Energy and Forests Appendix</b>	<b>130</b>

## Acknowledgments

I could not have written this dissertation without a lot of help from many people. I am enormously grateful to Severin Borenstein and Lucas Davis for their time and guidance. I can't imagine a better way to learn than by working and talking with them. I am lucky to have been a part of the community that Severin has created at the Energy Institute at Haas. The seminars and lunch table discussions taught me a lot of economics, but they also taught me how research can contribute to the world outside of academia. Max Auffhammer, Meredith Fowlie, and Catherine Wolfram provided generous advising and insight. Karen Notsund and Casey Hennig helped me in innumerable ways.

The ideas in this dissertation were all improved through conversations with my office-mates, especially Walter Graf and Erica Myers. Walter, Angeli Kirk, Carl Nadler, and Morgan Levy supported me from the beginning of our graduate school experience. Paula Pedro made sure I got to work on time. I also received generous financial support from the National Science Foundation Graduate Research Fellowship Program, the Energy Institute at Haas, and the Graduate Division.

# Chapter 1

## Drilling Like There's No Tomorrow: Bankruptcy, Insurance, and Environmental Risk

### 1.1 Introduction

In almost all modern legal systems the debts of insolvent parties can be eliminated through bankruptcy. Bankruptcy protection benefits society by improving insolvent actors' work incentives and by mitigating coordination problems among creditors.<sup>1</sup> However, bankruptcy protection also distorts behavior by insulating actors from worst-case outcomes. For example, financial firms may become excessively leveraged, consumers may accumulate excessive personal debt, and governments may commit to unsustainable levels of public spending. A range of private contract features and public policies exist to combat this incentive problem, such as reserve requirements for banks, loan limits for consumers, and credit ratings for governments. These policies in turn affect many aspects of the economy.

One important implication of bankruptcy protection is that firms in hazardous industries will take excessive environmental and public health risks. In many countries, including the United States, accident liabilities and regulatory judgments can be discharged in bankruptcy. This limits firms' liability to their current assets, and thus decreases the safety incentives of firms with assets that are less than their worst-case liabilities. Economists call this the "judgment-proof problem" (Shavell, 1986). In addition to distorting short-run incentives, the judgment-proof problem may increase the share of small firms in the long run. Staying small allows firms to avoid the consequences of accidents, and therefore the need to invest in costly safety measures. The implication of the judgment-proof problem for society is too many harmful accidents. In addition, if there are economies of scale in production, the judgment-proof problem can inefficiently limit productivity.

This paper examines the effect of bankruptcy protection on market structure and envi-

---

<sup>1</sup>For a review of the economics of bankruptcy law, see White (2007).



ronmental outcomes in the onshore oil and gas industry. This industry is an ideal setting for this analysis. Extracting crude oil and natural gas involves a risk of severe environmental and health damages through water pollution, toxic gas releases, and explosions. The industry in the United States includes many small firms. In 2012, almost 5,000 firms reported production in Texas alone (the largest oil- and gas-producing state, and the setting for my analysis). The vast majority of these firms had less than two million dollars in annual revenue.

In order to credibly measure the market structure impacts of the judgment-proof problem, the empirical analysis exploits quasi-experimental variation in the required surety bond amounts for oil and gas producers in Texas. Surety bonds are insurance contracts that obligate the insurer to pay the state for environmental costs left behind by bankrupt oil and gas producers. Bond mandates cause firms to internalize accident risk through the premiums they pay to the insurer. Firms with poor safety records and financially weak firms with little incentive to exercise environmental care will face high premiums. Surety bonds are complicated by moral hazard, as I discuss later, but in general a bond requirement improves firms' safety incentives.

Texas introduced bond requirements for some oil and gas producers in 1991, and for all producers in 2001. For both policy changes, firms were required to comply with the new rules by the date of their annual operating license renewal. These license renewal dates are determined by the anniversary of a firm's first license application, and are thus distributed throughout the year. This quasi-random variation in timing allows me to cleanly identify the effects of the policy changes. The analysis relies on a novel dataset of firm entry and exit, oil and gas production, and environmental outcomes for the universe of oil and gas producers in Texas, created by merging several different administrative datasets. Texas is the largest oil- and gas-producing state in the U.S. and keeps detailed records on production and environmental outcomes. Thus, it is an ideal geographic setting.

I find that greater internalization of environmental costs changed the industry structure. The 2001 universal bond mandate pushed about 6% of producers out of the market immediately. This was almost a 70% increase from the normal background rate of exit. The exiting firms were primarily small and were more likely to have poor environmental records. Finally, among firms that remained in business, the bond requirement reduced oil production among the smallest 80% of firms by about 4% on average, while production by the largest 20% of firms was unaffected. These results suggest that the ability to easily avoid environmental responsibilities prior to bonding inflated the number of small firms and their production.

Environmental outcomes also improved. After the universal bond requirement, the number of firms leaving their wells unplugged at the end of production (which creates a serious risk of groundwater pollution) decreased sharply. Well blowouts and violations of water protection rules also decreased substantially. These results suggest that by screening out firms with the least incentive to take care and increasing accountability for firms that remained in business, the bond requirement mitigated the harmful incentive effects created by bankruptcy protection.

The results of this study confirm the central predictions of a theoretical literature that previously has had little high-quality empirical validation. Existing empirical studies of liability

and market structure are based on cross-sectional comparisons across states, industries, or firm sizes. This study departs from previous work by providing credible quasi-experimental evidence that the judgment-proof problem is an important determinant of industry structure and environmental outcomes in hazardous industries. In addition, the use of firm-level administrative microdata allows for much more detailed analysis than has been possible in previous work. For example, I am able to measure how firm-level output responds to insurance requirements. The results bolster concerns among economists and policymakers about judgment-proof issues in other sectors, including landfills, underground storage tanks, small-scale manufacturing, and hazardous materials transportation.

This study also makes a theoretical contribution by extending existing models of the judgment-proof problem to allow firms to vary in output. This formalizes the relationship between bankruptcy and industry structure. The model yields clear, testable predictions about the effects of policy changes that mitigate the incentive problems created by bankruptcy, like the bond requirements that I observe in the empirical analysis.

Finally, this study contributes to our understanding of safety regulation in one of the most important industries in the world. Hydraulic fracturing has led to meteoric increases in oil and gas development in the United States in the past ten years. More than 15.3 million Americans have had an oil or gas well drilled within one mile of their home since 2000.<sup>2</sup> The shale boom has motivated a great deal of empirical research on outcomes of energy development such as environmental or wealth impacts (Allcott and Keniston, 2014; Darrah et al., 2014; Muehlenbachs, Spiller, and Timmins, 2014). This study targets a different question that has not been as widely explored: Is society, through regulation, successful in inducing oil and gas producers to balance profits and environmental risk in a socially efficient way? This study demonstrates one important reason that some firms depart from socially desirable behavior. The results also suggest that there would be benefits from increasing bond requirements in oil- and gas-producing states to at least the level required in Texas, and possibly further.

The rest of the paper is organized as follows: The following section discusses liability, bankruptcy, and market structure. Section 3 proposes a model for how bankruptcy protection affects firm size, output, and safety effort. Section 4 discusses the oil and gas industry, the empirical strategy, and the data. Sections 5 and 6 discuss the results. Section 7 concludes.

## 1.2 Background

### Liability and Industry Structure

The classic Shavell (1986) model of the judgment-proof problem shows that individuals whose potential liability exceeds their assets take inadequate care to prevent accidents (*e.g.*, they drive their cars recklessly) and engage too often in activities that may harm others

---

<sup>2</sup>Gold, Russell and Tom McGinty. “Energy Boom Puts Wells in America’s Backyards.” *Wall Street Journal*. October 25, 2013.

(*e.g.*, they drive too much). The same reasoning applies to firms in hazardous industries. Because safety effort is costly, firms that cannot be compelled to pay for accident damages will underinvest in accident prevention (Shavell, 2002). As I show in Section 1.3, they will also produce too much.

In industries where accidents and safety effort are expensive, the ability to avoid liability creates a cost advantage for financially weak firms. This means that firms may seek to strategically limit their asset exposure. One simple strategy is to keep the firm small.<sup>3</sup> Small firms have few assets to be seized after accidents. They can also be quickly dissolved in anticipation of accident claims, as in Boyd and Ingberman (2003). And, as I propose in Section 1.3, the probability that total accident damages will exceed total assets is larger when firms have fewer projects. Limiting firm size limits liability exposure, at the expense of economies of scale.

In the classic model of long-run competitive equilibrium, firm size is given by the unique output level that minimizes the U-shaped long-run average cost (LRAC) function (Viner, 1932). However, a large body of empirical research has documented disparities in firm size that are inconsistent with the Viner model (Bain, 1956; Bloom, Sadun, and Van Reenen, 2012). More recent models seek to explain firm size more directly. In Lucas (1978), firms differ in “managerial technology.” Better-managed firms are more productive and grow larger than poorly-managed firms, but span-of-control problems limit the growth of even the best-managed firms. In Banerjee and Duflo (2005), imperfect capital markets prevent all firms from adopting a capital-intensive technology with low per-unit production costs. The “missing middle” literature in development economics asks whether imposing regulations only on large firms keeps firms small (Rauch, 1991; Hsieh and Olken, 2014).

In industries with significant liability risk, bankruptcy protection may also affect firm size by creating an incentive to stay small. Surprisingly, formal models of the judgment-proof problem have had little to say about the choice of firm size. For example, Shavell (2002) assumes that firms produce only a single unit, and Pitchford (1995) describes firms as considering a single risky project.<sup>4</sup>

A few empirical studies have examined the judgment-proof problem. Alberini and Austin (2002) compares toxic releases in states with strict liability rules vs. negligence standards. They find that accidents are less likely in states with strict liability, except among the smallest firms. Ringleb and Wiggins (1990) argues that the number of small firms in dangerous industries in the United States increased during the 1970’s, at the same time that enforcement of liability for latent hazards like workplace carcinogen exposure became more aggressive. These and other existing empirical studies make cross-sectional comparisons across jurisdictions, firm sizes, or industries. These types of comparisons are vulnerable to omitted variables bias. For example, firm size may be correlated with experience or skill. In addition,

---

<sup>3</sup>Other strategies include contracting out risky activities to small firms (Brooks, 2002); premature dissolution (Boyd and Ingberman, 2003); and financing with securitized debt that is senior to accident claims (Che and Spier, 2008). See LoPucki (1996) for a general review of these strategies.

<sup>4</sup>Similarly, Ganuza and Gomez (2011) and Che and Spier (2008) assume identical output across firms while allowing for strategic choice of asset level and capital structure, respectively.

data on firms and accidents are often only available at highly aggregated levels, which limits the detail of many analyses. There is an opportunity for highly credible quasi-experimental approaches using detailed, firm-level data to provide evidence on how bankruptcy protection affects industry structure and safety effort.

## Effects of An Insurance Mandate

One widely-used policy to mitigate the judgment-proof problem is to require firms to have liability insurance or bonds.<sup>5</sup> Firms comply with bond requirements either by depositing valuable assets with the regulator, to be returned once production is completed safely; or by purchasing a surety bond from an insurer. Surety bonds are a promise by the insurer to pay the state up to the value of the bond if the insured firm goes out of business and leaves some environmental cost.<sup>6</sup> Firms with few assets typically choose surety bonds instead of cash bonds.

Insurance and bond requirements are commonplace. Most U.S. states require taxi firms to purchase liability insurance or surety bonds. Recently, the lack of a similar insurance requirement for transportation network companies like Uber became a public policy issue.<sup>7</sup> In construction, contractors must purchase bonds that will pay for project completion if they go out of business. Bonds or liability insurance are also required for owners of landfills and underground chemical storage tanks, both of which can cause serious pollution problems. In oil and gas extraction, all of the major oil- and gas-producing states and the federal Bureau of Land Management require bonds, although in many cases they are very small relative to potential damages.

Insurance and bond requirements mitigate the judgment-proof problem because firms internalize accident costs through premiums. Depending on how well insurers can observe safety effort, these policies have both intensive- and extensive-margin benefits. If effort is observable, then premiums will be conditioned on effort, the firm internalizes the full expected costs of accidents, and safety effort and industry structure will both be socially optimal. In the opposite extreme where safety effort is completely unobservable, the benefits are all along the extensive margin. Mandated insurance changes industry structure but not safety effort. Firms exercise the same safety effort as in the absence of the requirement,<sup>8</sup> and pay insurance premiums that reflect expected damages contingent on this sub-optimal level of safety effort. Insurance premiums act like a tax that discourages participation in the

---

<sup>5</sup>Other responses include direct regulation of safety (Shavell, 1984) and the extension of liability to firms' business partners (Kornhauser, 1982; Pitchford, 1995; Boyd and Ingberman, 2003).

<sup>6</sup>Bond requirements often allow firms to purchase surety bonds (from insurers) or irrevocable letters of credit (from banks). These instruments are very similar, and I follow the convention established by regulators of referring to both as "surety bonds" or "bonds".

<sup>7</sup>Dolan, Christopher. "Viewpoints: Uber should have enough insurance, just like anyone else". *Sacramento Bee*. August 8, 2014.

<sup>8</sup>In fact, Shavell (2005) shows that safety effort decreases since insurance premiums reduce the firm's assets and thus its liability exposure.

activity, screening out some firms whose revenues before the policy exceeded private but not social costs (Polborn, 1998).

Using state-level panel data on annual releases from underground storage tanks, Yin, Pfaff, and Kunreuther (2011) finds decreased leaks in seven states after tank owners were required to purchase private liability insurance. This cross-sectional evidence suggests that insurance mandates can improve environmental outcomes. However, the highly aggregated data limits conclusions about firm-level responses or industry composition.

### 1.3 Model

This section proposes a model of how bankruptcy protection leads small firms to undervalue safety, and leads to too many small firms. It extends the Shavell (1986) model to allow for endogenous selection of firm size and output. Previous models of judgment-proof firms treat output or the number of projects as constant across firms, which ignores the stylized fact that small producers are more likely than large producers to be judgment-proof.<sup>9</sup> I use the model to show how the judgment-proof problem inefficiently raises the number of small producers, and leads each of the judgment-proof firms to produce more than is efficient while exerting less-than-efficient safety effort. I also show how bond requirements partially mitigate this market failure. In addition to formalizing the relationship between industry structure and liability, the model yields a clear set of theoretical predictions to take to the data.

#### Profit Maximization in a Hazardous Industry

A homogeneous good is produced by risk-neutral firms in a competitive industry. Production may cause accidents that harm others. When an accident occurs, the damages are a constant  $h$ . Firms can reduce the likelihood of accidents by exerting costly safety effort. The level of safety effort is a continuous variable  $x$ , and the cost is normalized to one, so that expenditures on safety are also  $x$ . The probability that an accident will occur for any given unit of production is  $\gamma(x)$ , which is decreasing in  $x$ . When an accident occurs, a fine equal to  $h$  is levied.

Accident risk and safety effort increase linearly per unit of output ( $q$ ). For example, each additional oil well or manufacturing plant presents a hazard and requires independent safety effort.<sup>10</sup> The function  $f(v; q, x)$  is the probability distribution of total accident damages  $v$  for a firm. Damages are bounded by  $hq$  (*i.e.*, an accident with every unit of production). I assume that accidents are independent, so  $E[v] = \gamma(x)qh$ .

A firm chooses  $q$  and  $x$  to maximize the profit function,

$$pq - c(q) - xq - \gamma(x)qh \tag{1.1}$$

---

<sup>9</sup>An exception that I became aware of after developing the model in this paper is Van 't Veld (2006).

<sup>10</sup>A more complex model could also include fixed costs of safety. The primary implication would be that only firms above some minimum efficient scale would find it profitable to invest in safety, reinforcing the judgment proof problem.

Profit-maximizing safety effort  $x^*$  minimizes effort costs plus fines:  $q[1 + \gamma'(x^*)h] = 0$ . At  $x^*$ , the reduction in expected fines from increased safety effort equals the marginal cost, which is one. Profit-maximizing output  $q^*$  equates marginal cost, including safety effort and expected fines, with price:  $p = c'(q^*) + x^* + \gamma(x^*)h$ . Firms internalize accident costs through fines, so  $(x^*, q^*)$  is also socially optimal.

## The Judgment-Proof Problem With Output Choice

This section introduces a parameter  $y$ , which is the assets owned by the firm that can be seized to pay fines. If  $y$  is less than damages, the difference is discharged in bankruptcy. To begin, assume  $y$  is given exogenously. The profit function becomes,

$$pq - c(q) - xq - \begin{cases} \gamma(x)qh & y \geq hq \\ \left[ \int_0^y vf(v; q, x)dv + \int_y^{hq} yf(v; q, x)dv \right] & y < hq \end{cases} \quad (1.2)$$

When  $y$  is greater than or equal to maximum possible liability, the profit function is unchanged. However, when  $y$  is less than maximum possible liability, the firm will not consider the full expected costs of accidents. Instead, it considers the mean of a truncated damage distribution, valuing damage outcomes greater than  $y$  at  $y$ . Expected damages are replaced in the profit function by the probability-weighted sum of damages from zero to  $y$ , plus  $y$  for all larger damage outcomes.<sup>11</sup>

A firm with  $y < hq$  is, to some degree, judgment-proof. For at least some realizations of total damages, the firm's assets will be less than the damages. It chooses safety effort  $x(y) < x^*$ , increasing the risk of accidents. The intuition is simple. The benefit of accident prevention is smaller for a judgment-proof firm than a responsible firm because the judgment-proof firm does not internalize the full expected cost of accidents.

Judgment-proof firms also choose  $q(y) > q^*$ . Again, the reasoning is straightforward. If the firm fully internalized accident costs, the safety cost of another unit of output would be the cost of socially optimal safety effort plus expected accident damages at that effort level:  $x^* + \gamma(x^*)h$ . For a judgment-proof firm starting from the same level of output, the safety cost of another unit of output is the cost of suboptimal safety effort plus the change in the expected private cost of total accident damages. This is  $x(y) + A$ , where  $A = \frac{d}{dq}[\int_0^y vf(v; q, x)dv + \int_y^{hq} yf(v; q, x)dv]$ . It was already established that  $x(y) < x^*$ . It is also clear that  $A < \gamma(x^*)h$ , because for damage realizations greater than  $y$ , the firm will not bear the additional damages imposed by this unit; and both firms face identical damages when damages are less than  $y$ .<sup>12</sup> So, marginal private safety costs are smaller when firms are judgment-proof, leading to higher production compared to a same-sized firm that was responsible for all damages.

<sup>11</sup>When accidents are independent, the number of accidents follows a binomial distribution, and expected fines for the judgment-proof firm can be expressed more precisely as,

$\sum_{k=0}^y kh\phi(k) + \sum_{k=y/h}^q y\phi(k)$  where  $\phi(k)$  is the binomial pmf,  $\binom{q}{k}\gamma(x)^k(1 - \gamma(x))^{q-k}$ .

<sup>12</sup>The privately-optimal  $x$  for the judgment-proof firm may also increase with  $q$ . However, re-optimizing  $x$  can only reduce marginal cost relative to the old  $x$ .

## The Judgement-Proof Problem and the Size Distribution of Firms

This section shows how the judgment-proof problem increases the number of small firms. This requires some additional assumptions. I assume heterogeneity between firms in technology, represented by  $\theta$ . A firm's production cost for quantity  $q$ , exclusive of safety effort and fines, is given by  $c(q; \theta)$ . Firms with larger  $\theta$  have higher fixed costs, but lower minimum average cost. This feature of the model can be motivated in several ways, as discussed in Section 1.2. For example, in the Lucas (1978) model of firm size, larger  $\theta$  would correspond to more skilled management.

Secondly, I assume that small firms are more likely to have assets less than total damages. This assumption is widely made in the accident economics literature, but is not often tied to a formal economic model. One way to motivate this feature is to assume that  $y$  increases linearly in  $q$  (for example, because every well or plant adds valuable physical assets). In that case, the volatility of total damages relative to mean damages (*i.e.*, the coefficient of variation) decreases in  $q$ . A firm with many plants is less likely to have accidents at all of its plants than a firm with one plant, when accidents are independent. So, the probability that accidents will bankrupt larger firms is lower because pooling projects reduces the variability of total damages.<sup>13</sup> A second motivation for this assumption is that it is easier for small firms to pursue a “fly-by-night” strategy of premature dissolution, as in Boyd and Ingberman (2003). It is less costly for firms with few assets to quickly strip value out of the firm in anticipation of liability claims for accident damages.

Figure 1.1 shows the potential for inefficient market structure. The thick solid curves show long-run average cost curves for two types of firms with different  $\theta$ , when firms fully internalize expected accident damages. These curves represent the vertical sum of average production costs,  $ac(q; \theta)$ , and average safety costs (effort plus expected damages),  $x^* + \gamma(x^*)h$ . Because  $\theta_2 > \theta_1$ , the minimum average cost for Type 2 firms is below and to the right of that of Type 1. When firms internalize accident damages, large firms are most privately efficient and will dominate in a competitive market setting.

The dashed curves show how bankruptcy lowers private average costs for small firms. When  $q$  is small, the firm is unlikely to pay the full costs of accident damages because damages are likely to exceed assets. It chooses a low level of safety effort, and its overall safety costs, shown by the dashed safety cost curve at the bottom of the figure, are low. As  $q$  increases, the share of expected damages internalized by the firm increases, so that per-unit safety costs approach  $x^* + \gamma(x^*)h$ . The dashed average cost curve shows overall average cost for Type 1 firms when bankruptcy limits liability. Now small firms are the most privately efficient and will dominate in a competitive market setting.<sup>14</sup>

<sup>13</sup>Technically, this requires the additional assumption that  $y$  is at least as large as expected per-unit damages under socially optimal safety effort. If accidents are rare, this assumption is not restrictive.

<sup>14</sup>In this simple model, a single type of firm will dominate the market in equilibrium. In reality, without infinite potential entry of each type, there would be a range of firm sizes in the market. However, which type was most privately efficient would still be determined by the reasoning here, so that bankruptcy protection would increase the number of small firms.

This model demonstrates a tradeoff between economies of scale and the ability to avoid accident costs. For Type 2 firms, economies of scale are large. These firms maximize profit by producing a large amount of output, even though this large firm size means they must fully internalize the expected costs of accidents. Type 1 firms are small enough to avoid some accident costs, but economies of scale are limited. These types are two examples from a range of possible firm sizes that depends on the management or other technologies available in the market. The privately optimal firm size balances avoided liability and economies of scale.

What are the welfare effects of the judgment-proof problem? The per-unit social cost of the judgment-proof firm's production is higher than under full cost internalization. This is because, by definition, safety effort  $x^*$  minimizes the sum of effort costs and accident costs. Any other level of safety effort results in higher total social costs per unit of output. The marginal unit of production by judgment-proof firms has negative net benefits due to suboptimal safety effort and excessive production. In addition to this externality problem, if output produced by judgment-proof firms would otherwise have been produced by large firms, there are also foregone economies of scale.

## Alternative Assumptions

It is worth briefly considering some of the assumptions in the previous sections. The model ignores other well-known problems with liability regulation. For example, it assumes that accidents are always detected and that penalties exactly equal social damages. A complete model of liability regulation would include these features and several others described in Shavell (2007) and related works. I have abstracted away from these issues to focus on the judgment-proof problem. Adding these features to the model would not fundamentally change the predictions.

The model also treats firms as risk-neutral, so that there is no private demand for liability insurance. If firms are risk-averse, this may not be true. However, note that firms only have incentive to purchase insurance for losses that they would otherwise pay. Judgment-proof producers have no incentive to purchase coverage in excess of their assets, so the existence of a private insurance market is insufficient to eliminate the judgment-proof problem.

The preceding section also does not address capital structure. One implicit assumption is that firms cannot issue debt that is senior in repayment to accident damages. If this were possible, firms of any size could eliminate liability exposure by issuing debt secured by all of the firm's assets, as in Che and Spier (2008). In the event of an accident, all assets would already be pledged to senior creditors. For environmental damages in the United States, the assumption that this is impossible mirrors reality. Several federal and state laws make it difficult for lenders to foreclose on assets involved in environmental incidents, effectively subordinating a secured creditor's claim to environmental costs.<sup>15</sup> In the specific case of

---

<sup>15</sup>Interpretations of the law have varied, but, for example, a creditor who takes ownership of a Superfund site through foreclosure faces a non-trivial legal risk of being held liable as an owner (Harkins, 1994; Murray



Texas oil and gas extraction, for certain types of clean-up costs the state has a lien against insolvent producers' assets that is senior to secured debt.<sup>16</sup>

When the firm is financed with debt, the owner's incentives to operate safely are reduced because some of the losses due to accidents will be borne by creditors. This raises the question of why firm owners in hazardous industries do not seek to become highly leveraged in order to fully externalize the potential loss of  $y$ . The answer is that, as long as debt is junior to accident damages, borrowing costs for a firm with this strategy would be very high because lenders would fear losses.

Finally, for the application to oil and gas extraction, it is worth considering how to treat site reclamation in the model. Plugging wells and remediating waste storage pits is an important investment of safety effort that reduces groundwater contamination. In practice, regulators do not only use a liability rule to regulate reclamation. Instead, reclamation is mandated at all sites. For this element of safety effort, penalties to the firm are a deterministic function of effort. Firms pay fines with probability 0 if they remediate and probability 1 if they do not. What matters is that small firms can avoid both liability-based penalties and effort-based penalties through bankruptcy.

## Insurance Requirements Mitigate the Judgment-Proof Problem

This section shows how a surety bond requirement leads small producers to internalize a larger share of damages. Bonding increases care and reduces participation by judgment-proof firms. I also discuss how setting the bond amount too high can limit participation by firms whose operation would be welfare-improving.

Firms purchase surety bonds from insurers. These are contracts that obligate the insurer to pay the state any positive difference between the firm's damages and its assets. Maximum payments by the surety are limited by the face value of the required bond,  $\beta$ , which is chosen by the state. In a competitive market, insurers sell surety bonds at a price  $\pi$  that just covers expected losses plus underwriting expenses.

$$\pi = \int_y^{\beta+y} (w - l)f(v)dv + \int_{\beta+y}^{hq} (\beta - l)f(v)dv + u \quad (1.3)$$

The amount by which damages exceed the firm's assets is  $w = \max[0, v(q; x) - y]$ . Collateral required by the insurer is  $l$ . If damages are less than  $y$ , the insurer pays nothing. The first integral term represents the insurer's losses when damages are between  $y$  and  $\beta + y$ , so that the insurer pays the full difference between damages and assets. The second integral represents the insurer's losses when damages exceed  $\beta + y$ , so that the insurer pays  $\beta$  (so, unless the bond is set at or above worst-case damages, insurers do not fully internalize expected damages). Underwriting expenses are  $u$ .

---

and Franco, 2011).

<sup>16</sup>Texas Natural Resources Code Section 89.083.

A surety bond does not directly increase a firm’s liability exposure. Damages greater than  $y$  are paid by the insurer, not the firm, so that there is moral hazard on the part of the firm. However, the insurer will screen clients and design contracts to limit losses. Insurers can directly monitor safety effort, adjust rates based on assets or accident history, use credit reporting to punish firms who default, and/or demand collateral (sometimes from outside the business). These measures will induce firms to expend some safety effort, although this level of effort  $\hat{x}$  is likely to be below  $x^*$ . Since  $\hat{x} < x^*$ , the equilibrium surety bond price will reflect a higher level of expected damages than under  $x^*$ .

The surety bond requirement also introduces transaction costs, which are borne by the firm. I assume that collateral is invested at a rate lower than the firm’s opportunity cost of capital, with the proceeds accruing to the firm. Collateral requirements cost the firm  $lr$ , where  $r$  is the difference between the firm’s opportunity cost of capital and the rate at which collateral is invested. Underwriting costs  $u$  are also borne by firms.

For a bond amount equal to  $\beta$ , a small firm’s expected profit is,

$$p\hat{q} - c(\hat{q}; \theta) - \hat{x}\hat{q} - \int_0^y v f(v)dv - \int_y^{h\hat{q}} y f(v)dv - lr - \pi(\hat{x}, \hat{q}, y) \tag{1.4}$$

The bond requirement causes intensive and extensive margin improvements. On the intensive margin, the bonded firm invests higher safety effort ( $\hat{x}$ ) and produces less ( $\hat{q}$ ) than the unbonded firm because surety premiums improve incentives. On the extensive margin, firms with few recoverable assets will face high bond premiums, since insurers know it is not in those firms’ private interest to exercise adequate safety effort. This leads to a desirable change in industry structure as financially weak producers are screened out.

A bond requirement may also exclude firms whose operations would be socially efficient. Underwriting and collateral costs increase bond prices above expected losses, potentially excluding some firms that would be barely profitable with full cost internalization and no transaction costs. Other potential inefficiencies in insurance markets could also result in coverage not being available to responsible firms. I return to this topic in Section 1.6.

## Testable Predictions

This model makes several testable predictions about policy changes that reduce the ability to avoid bankruptcy in industries with large potential liabilities. There should be exit by small producers, as firms that had been privately profitable but not socially efficient are pushed out of the market. These exiters should have had more accidents prior to the policy change than firms that stay in the industry. Small firms that remain in the market should decrease their output as they internalize a larger share of marginal safety costs. Finally, environmental incidents should decrease due to increased safety effort per unit of production.

## 1.4 Empirical Analysis and Data

The setting for the empirical analysis is the onshore oil and gas extraction industry in Texas. This is an excellent industry in which to study liability and bankruptcy. There are significant environmental risks and many small producers. Firms produce homogeneous products, and the industry is essentially perfectly competitive due to extremely liquid international markets. Finally, the size of the industry means that environmental incidents occur frequently enough to allow for detailed empirical analysis. Texas is the largest producer among the United States and keeps detailed data on production and accidents, making it an ideal geographic setting.

### Environmental Protection in Oil & Gas Extraction

Oil and gas extraction poses potential risks of water pollution, methane leaks, and releases of toxic gases and radioactive materials. Historically, water pollution has been considered the most serious risk. Crude oil, drilling chemicals, and saltwater produced along with oil and gas can all have severe health and natural resource impacts.

Extracting crude oil and natural gas involves drilling and production phases. Wells are drilled into underground formations thought to contain hydrocarbons. Successful wells will produce for twenty years or more, with declining production over time. Water contamination is a serious concern during both drilling and production.

During drilling, a cement and steel well casing is constructed around the well bore. Correct construction of this casing is important to prevent leaks into groundwater. In addition, open pits for temporary storage of drilling wastes must be carefully sited and constructed. Finally, if drillers lose control of well pressure, there may be a “blowout”, possibly spreading oil or drilling fluids over a wide area. Blowouts can also cause fires, equipment losses, and injuries or deaths. If a blowout damages the underground resource, nearby mineral owners can claim damages. The total costs of a well blowout can be tens of millions of dollars (Jones, 2003).<sup>17</sup>

During production, crude oil or liquid wastes can leak from storage tanks, pits, pipelines, or trucks. In addition, degraded well casings can develop leaks into groundwater. Producers must monitor the integrity of aging wells. Once wells are no longer producing (or leaks are detected), they should be plugged with cement below groundwater depth. Plugging wells costs thousands of dollars per well, so operators have an incentive to avoid or delay it. In the extreme, wells may remain unplugged after the firm is dissolved. These “orphan wells” must be plugged with public funds. Because operators may hold inactive wells for years

---

<sup>17</sup>Because a firm’s demand for drilling services varies greatly in time and space, drilling is typically carried out by specialized contractors. Well owners and drillers negotiate the legal responsibility for accidents as part of their contracts. Typical contracts, such as the model contracts provided by the International Association of Drilling Contractors, assign most liability to well owners. Ultimately, both firms face some liability exposure, since the indemnities in these contracts are imperfectly enforceable. For example, either firm could become a target for liability claims if the counterparty became insolvent (Anderson, 1989; Jones, 2003).

before exiting, and because state agencies have limited budgets, state-funded plugging may come too late to prevent groundwater contamination. Between 1983 and 2008, orphan wells caused 17% of detected oil and gas groundwater contamination incidents in Texas, and 22% in Ohio (Kell, 2011).<sup>18</sup>

## Changes in Insurance Requirements

Credibly measuring the effects of the judgment-proof problem requires variation in firms' ability to avoid liability that is exogenous with respect to unobserved determinants of the outcome of interest. This is a difficult empirical challenge. A randomized control trial assigning asset levels or bond requirements to firms is unlikely to prove feasible. Cross-sectional comparisons of large vs. small firms, or across industries or jurisdictions, are highly vulnerable to omitted variables bias. For example, deep-pocketed producers may be more experienced or more skilled than financially weak producers.

This paper exploits variation over time in bond requirements. Stricter bond requirements reduce small firms' ability to escape environmental costs. Texas introduced a bond requirement for some oil and gas producers in 1991 and extended it to all producers in 2001. Both of these policy changes became binding on firms at the time of their first annual operating license renewal after the change. License renewal dates depend on the anniversary of the firm's creation, creating a quasi-random rollout. As I describe in detail in Section 1.5, this variation allows me to separate the effects of the policy changes from other time-varying determinants of industry composition.

Senate Bill 1103 in 1991 first introduced the bond requirement, but it required bonds for only some producers. Firms with an acceptable compliance history could instead pay a small annual fee called the "Good Guy Fee."<sup>19</sup> This option was chosen in 76% of license renewals during 1992–2001. In 2001, Senate Bill 310 extended the bond requirement to all oil and gas producers, eliminating the Good Guy Fee. This rule passed the legislature in June, 2001 and took effect in January, 2002. I focus the analysis primarily on the 2001 change because prior to this, few firms were bonded.<sup>20</sup>

Under both the partial and the universal bond requirements, the bond amount depended on the number and depth of wells. The formula was \$2 per foot of well depth across all wells. The mean depth of existing wells in 2001 was 3,300 feet.<sup>21</sup> Alternatively, producers

---

<sup>18</sup>For Texas, 17% of cases came from orphan wells; 33% from production activities; 43% from waste management and disposal; and 6% from drilling and well completion. For Ohio, 22% of cases came from orphan wells; 21% from production activities; 14% from waste management and disposal; and 40% from drilling and well completion.

<sup>19</sup>Firms with a four-year record of compliance with Commission regulations could pay the \$100 per year Good Guy Fee. Firms ineligible for the Good Guy Fee could also avoid bonding through an annual cash fee equal to 3% of the required bond amount. For unbonded operators, the 1991 law also introduced a \$100 annual fee for each inactive, unplugged well.

<sup>20</sup>S.B. 310 also allowed one final non-bond alternative, a sharply elevated annual fee equal to 12.5% of the required bond amount, to persist as a temporary alternative until 2004.

<sup>21</sup>This is the total depth of wells operated in 2001 divided by the number of wells.

could cover a number of wells with a “blanket bond.” Up to 10 wells could be covered with a \$25,000 blanket bond; 11 to 99 wells with a \$50,000 blanket bond; and over 100 wells with a \$250,000 blanket bond.<sup>22</sup>

Ninety-seven percent of bonded producers chose to purchase a bond from an insurer instead of posting their own assets as a cash bond.<sup>23</sup> The annual premium for a surety bond is typically 1-2.5% of the face value of the bond. However, for firms that are deemed to be high-risk because they are financially weak or have a poor safety record, premiums can exceed 10-15%. Bond issuers may also require collateral from firms deemed to be high risk (Gerard, 2000; Boyd, 2002; Kaiser and Snyder, 2009; Gerard and Wilson, 2009). Thus, the primary effect of the bond requirement was to substantially increase the operating costs of firms for which the market perceived a high risk of insolvency and environmental damage. A low-risk operator with five wells might pay \$330 per year in bond premiums, while a high-risk operator with the same number of wells might pay \$2,475 per year while also facing collateral requirements.<sup>24</sup>

The importance of the bond requirement to producers’ costs was reflected in the widespread news coverage it received. Bonding was controversial because of a perception that it pushed out small firms. An op-ed in the *Midland Reporter Telegram* stated, “Bonding is no problem for major oil companies and large publicly owned independent companies. The small independent, however, is finding bonding very difficult at best... If the Railroad Commission persists in its current bonding requirements it could put thousands of honest hardworking Mom and Pop operators out of business.”<sup>25</sup> Under the headline, “Can’t Afford the Bond? Then Don’t Run a Well,” the *San Antonio Express News* editorialized, “Texas is better off if only companies that can afford to be responsible environmental stewards stay in the oil and gas business.”<sup>26</sup>

The policy was primarily intended to reduce orphan wells, but the required bonds were conditioned on preventing all types of water pollution. The bond states that “all oil and gas activities and operations shall be carried out so as to prevent pollution of any ground or surface water in the state.”<sup>27</sup> Unlike liability insurance, if the state makes a claim against the bond, the insurer will seek repayment from the oil and gas producer. Thus, these bonds

---

<sup>22</sup>Between steps in the blanket bond schedule – at 10 to 11 wells and 99 to 100 wells – the marginal increase in the required coverage level is zero. It is worth considering whether this introduced new economies of scale. In practice, any effect was likely small. First, adding wells increased the surety’s risk exposure and so should have increased premiums for a given level of coverage. Second, bond costs were small for low-risk firms. Ignoring premium increases, a large firm acquiring a low-risk five-well firm would have saved about \$165 per year for each of the firms’ owners, following the example calculation in the following paragraph. These savings were likely small compared to other benefits and costs of merging. For high-risk firms, the benefits of combining with a large producer with a good reputation were higher. But the main reason for the savings was not the blanket bond; it was that the larger firm was a better insurance risk.

<sup>23</sup>This is the share of bonded license renewals from 2002-2005 with surety bonds or letters of credit.

<sup>24</sup>Low risk firm: 3,300 feet \* 5 wells \* 2%. High-risk firm: 3,300 feet \* 5 wells \* 15%.

<sup>25</sup>“PBPA members detailing problems getting bonds.” *Midland Reporter Telegram*. March 31, 2002.

<sup>26</sup>“Can’t Afford the Bond? Then Don’t Run a Well.” *San Antonio Express News*. August 9, 2002.

<sup>27</sup>Railroad Commission of Texas Blanket Performance Bond P-5PB(2). <http://www.rrc.state.tx.us>

transfer default risk from the state to the surety.

## Data

For this analysis, I construct a novel dataset on market structure and environmental outcomes. The core of the analysis relies on several administrative datasets from the Railroad Commission of Texas (RRC), the agency that regulates oil and gas production. Operator entry and exit dates come from the RRC “Organization Report” dataset. All operators must file an organization report annually by the anniversary date of the firm’s first filing. I define a firm’s exit date as 365 days after its final organization report renewal, since the firm chose not to renew its operating license as of this date.

Bond data were obtained through a public records request to the RRC. This dataset includes, each year for every operator: the type of bond (surety bond, “Good Guy” option, etc.); the required bond amount; and the number and depth of wells.

Oil and gas production data come from the RRC Production Database Query (PDQ) dataset (for 1993–2010) and the RRC Final Oil and Gas Annuals (FOGA) dataset (for 1990–1992). Both datasets report monthly crude oil and natural gas production at the lease level.<sup>28</sup> A lease is a parcel of land on which the producer has negotiated the right to explore for oil and gas, and may contain multiple wells. Each lease-month observation is matched to an operator using unique operator identification numbers.

Drilling data come from the RRC “Drilling Permit Master and Trailer” dataset. That dataset identifies “spud-in” and well completion dates for every well drilled between 1991 and 2010. Spud-in is the date that drilling began, and well completion is the date that the well first produced oil or gas, which coincides with the conclusion of rig work.<sup>29</sup> Wells are matched to operators using the same operator identification numbers.

Environmental outcomes comes from several datasets. Orphan well data come from the RRC Orphan Well Database. This is a snapshot of orphan wells that have not yet been plugged by the state as of March 14, 2014.<sup>30</sup> Information on environmental rules violations comes from the RRC’s online Severance Query Database. I am primarily interested in violations of Statewide Rules 8 and 14. Statewide Rule 8 (“Water Protection”) governs water quality protection during drilling and production. Statewide Rule 14 (“Plugging”)

---

<sup>28</sup>I include casinghead gas (gas from wells that primarily produce oil) in natural gas production. I include condensate (a liquid petroleum product from gas wells) in crude oil production. In practice, condensate is a slightly different product and its market price can differ slightly. However, it is a small share of total production, and excluding condensate does not meaningfully affect the results.

<sup>29</sup>The data extend back to 1976; however, Kellogg (2011) reports that spud-in dates were not reliably recorded prior to 1991. Also following Kellogg (2011), I disregard completion dates for wells with reported drilling times less than 0 or greater than 180 days. For wells with no valid completion date, I assume that rig work took place for the average drilling time of 20 days.

<sup>30</sup>A comprehensive list of all wells that have ever been orphaned is not available from the RRC. Because I am primarily interested in the change in the rate of well orphaning with bond policy changes, an incomplete list does not affect my analysis as long as the state did not preferentially plug wells orphaned just before or just after rules changes.

requires that inactive wells be plugged promptly. Information on well blowouts comes from the RRC “Blowouts and Well Control Problems” list. Oil and gas operators are required to notify the Commission of all well blowouts. Orphan wells and rules violations are merged to operators using operator identification numbers. The blowout data do not include these numbers. I match blowouts to operators using unique lease identification numbers where provided. For the remaining blowouts, I carefully match on lease name and operator name.

The final dataset includes all operators in the state with positive oil or gas production between 1990 and 2010. This includes 17,672 firms.

## Descriptive Evidence of the Judgment-Proof Problem

Before proceeding to the empirical results, this section summarizes descriptive evidence that suggests the judgment-proof problem may have been important in this industry. In 2001, the year before the expanded bonding requirement took effect, 5,302 firms in Texas reported oil and gas production. Figure 1.2 shows the distribution of their revenues. Above \$15 million, the horizontal axis is truncated to show the five largest producers. The vertical axis shows the number of firms in each 1 million dollar bin. While the largest firms produced over one billion dollars worth of oil and gas, most firms had revenues below \$1 million. Table 1.1 lists the quintiles of the revenue distribution. The 80th percentile of the revenue distribution is only \$1.4 million dollars, and the 20th percentile is \$33,000.

Figure 1.3 shows that environmental incidents were concentrated among small operators. I calculate average annual production for all firms from 1990–2001. The x-axis shows the sum of annual production across all firms. Producers are ordered left to right along this axis from smallest to largest. The vertical axis shows the cumulative share of environmental incidents. Almost 100% of orphan wells, 95% of field rules violations, and 40% of well blowouts are associated with the 20% of total production that comes from the smallest firms.

Finally, there is also evidence of high bankruptcy rates. According to a 2003 report by the State Review of Oil and Natural Gas Environmental Regulations (STRONGER), during 2001 and 2002 (before the bond requirement was fully implemented), the state was unable to collect 68% of the penalties assessed for oil and gas rules violations. The most common reason that these fines were uncollectible was bankruptcy.<sup>31</sup>

## 1.5 Results

There are three sets of empirical analyses related to industry composition, firm-level output, and environmental outcomes. This section presents the results of each. As discussed in Section 1.4, I focus the empirical analysis primarily on the introduction of the universal bond requirement in 2001, because prior to this many producers were able to avoid bonding through non-bond alternatives like the “Good Guy Option.”

<sup>31</sup> “The RRC assessed \$5,183,832 in penalties and collected \$1,728,595... Penalties are most often uncollectible because the company has gone out of business and has no assets” (STRONGER, 2003).

## Industry Composition

The model in Section 1.3 predicts that increased bond requirements will cause some firms to exit because they are no longer profitable after internalizing a larger share of environmental costs. I measure the effect of the 2001 bond requirement on producer exit using a regression discontinuity (RD) design. This design takes advantage of the twelve-month bond rollout by comparing firms who renewed within a narrow window in time around the implementation of the policy.<sup>32</sup> I run the regression,

$$1[Exit]_{it} = \alpha + \beta_1 1[Begin\ Rollout]_t + \beta_2 1[End\ Rollout]_t + \beta_3 T_t + X_t \beta_4 + \psi_m + \eta_{it} \quad (1.5)$$

The dependent variable  $1[Exit]_{it}$  is an indicator variable equal to one if firm  $i$  exits in month  $t$ . I observe exit decisions once per year per firm, during the firm's renewal month. Thus, while the analysis is at the monthly level, there is one observation per firm per year. After firms exit they are removed from the sample.  $1[Begin\ Rollout]_t$  is an indicator variable equal to one in January 2002 and all later months.  $1[End\ Rollout]_t$  is an indicator variable equal to one in January 2003 and all later months.  $T_t$  is a local polynomial in the running variable, the number of months before or after January 2002.  $X_t$  includes monthly crude oil prices.<sup>33</sup>  $\psi_m$  is a set of month-of-year fixed effects. Controlling for output prices and month of year improves precision and allows comparison of the effect of the policy to the effects of short-term fluctuations in oil and gas prices. The constant term  $\alpha$  gives the background level of exit prior to the policy change.  $T_t$  is centered at January, 2002 and oil and gas prices are centered at their sample means, so  $\alpha$  gives the level of exit that would have been expected immediately prior to the policy change, under average oil and gas prices (and in January when month-of-year fixed effects are included).

The error term  $\eta_{it}$  includes unobserved determinants of  $1[Exit]_{it}$ . The identifying assumption in this analysis is that  $\eta_{it}$  does not change discontinuously at the implementation threshold. Under this assumption, the bond rollout allows for clean measurement of  $\beta_1$  by comparing exit during a small number of months after implementation to exit during a small number of months prior to implementation. As I show below, this identifying assumption is bolstered by the pattern of exit observed at the beginning and end of the bond rollouts in 1991 and 2002. There are clear, symmetric discontinuous changes at each of these thresholds. Thus, it is unlikely that the increased exit at the time of the policy change was solely driven by an unobserved idiosyncratic shock.

### Industry Composition: Graphical Results

Figure 1.4 shows the raw data on producer exit. The dots represent the number of firms leaving the market each month. Prior to the introduction of the partial bond requirement

<sup>32</sup>As suggested in Lee and Card (2008) for RD designs with discrete support, I cluster the standard errors according to the running variable. The running variable is month; I cluster by quarter.

<sup>33</sup>Crude oil and natural gas prices are highly correlated during this period (correlation coefficient 0.8). To simplify interpretation of the price effects, I include oil prices only. Appendix Table A.2 shows results controlling for oil and natural gas separately.



in September 1991, about 40 firms exited each month. With the bond requirement, the rate of exit increases sharply to over 100 firms per month and stays high for 12 months. After the 12 months, the level of exit decreases sharply. The same pattern of sharply increased exit for twelve months accompanies the implementation of the universal bond requirement in 2002. In 2002, the level of exit approximately doubles from about 60 to about 120 firms per month. Appendix Figure A.1 shows a similar figure for net entry (entry minus exit) by month. The pattern is very similar.

Figure 1.5 shows a graphical version of the RD estimator for the 2001 universal bond requirement. The sample includes 1997–2006. The x-axis shows months before and after January 2002. The dots are monthly means of the residuals from a regression of  $1[Exit]_{it}$  on a constant term, month-of-year fixed effects, and monthly oil prices. The fitted curves are a quadratic polynomial fit for 1997–2001, a separate linear fit for 2002, and a separate quadratic polynomial fit for 2003–2006. There is a clear increase of about five percentage points in the share of firms exiting in January, 2002. At the end of the twelve-month implementation period, there is a discontinuous decrease in exit to approximately the pre-implementation level.

Figure 1.6 shows the same figure, by quintile of firm-level average annual oil and gas production. These quintiles are defined based on average production during 1996–2007 in barrel of oil equivalents (BOE).<sup>34</sup> The effect of the universal bond requirement is largest for small firms. Among the smallest 20% of producers, there is a clear increase in exit of about 15 percentage points in January 2002. In the second quintile, the effect is about 10%; in the third, it is about 5%, and in the fourth it is about 4%. In each of these middle quintiles the effect is growing smaller, but is clearly visually distinguishable in every case. In contrast, in the top quintile of firm size there is no distinguishable change in the share of firms exiting in January 2002.

For robustness, Appendix Figure A.2 shows the number of firms with license renewal dates in each month of the sample. These dates are assigned by the RRC and cannot be manipulated by firms. The bond requirement was implemented according to firms' assigned renewal dates, so firms could not avoid it by submitting their renewal paperwork early. If such manipulation had been possible, it may have introduced systematic differences between firms up for renewal before and after implementation. As expected given the implementation rules, however, the number of firms up for renewal is smooth across the implementation threshold.

### Industry Composition: Estimation

Table 1.2 shows regression results for equation 1.5. As in Figure 1.5, each specification includes a quadratic polynomial in time for 1997–2001; a separate linear polynomial for 2002; and a separate quadratic polynomial for 2003–2007. In Column (1), the implementation of

---

<sup>34</sup>One BOE is one barrel of oil or 6,000 cubic feet of natural gas. One BOE represents the approximate energy content of a barrel of oil, and is a commonly used metric for combining oil and gas into a single measure. Results are similar if I use value of production instead of BOE; however, BOE is preferred because changes in oil and gas prices over time affect the production value measure.

the expanded bonding rule causes a discontinuous increase in exit of 6.8 percentage points. This means that an additional 6.8% of the firms scheduled to renew their license each month chose to leave. Over 12 months, this effect would decrease the total number of firms in the industry by 6.8%. The baseline rate of exit before the policy change, given by the constant term, is 9%. The specification in Column (2) controls for crude oil prices in each firm's assigned renewal month. This decreases the estimated effect of the bond policy slightly to 6.3 percentage points. The effects of price increases on exit are negative, as expected. Oil prices are centered at their mean value in the panel, so the constant term in Column (2) describes the baseline exit rate at average oil and gas prices. Finally, the specification in Column (3) also includes calendar month fixed effects. Including the fixed effects has little effect on the estimates, which suggests that there are not important systematic differences in firms across assigned license renewal months.

These results suggest that the bonding requirement caused 6% of the firms in the industry to exit immediately. This is about a 65% increase from the normal rate of exit. To further contextualize this effect, it is interesting to compare the exit due to the bond mandate with the amount of exit caused by month-to-month fluctuations in output prices. The spot price of crude oil is one of the best predictors of future oil prices, even compared to futures prices (Alquist and Kilian, 2010). Thus, by reducing expected future revenue for producers, output price decreases play a similar role to generic cost shocks. Taking literally the estimates in Column (3) of Table 1.2, a \$10 decrease in the crude oil price would lead to a 0.5 percentage point increase in exit probability across all firms. So, the immediate contraction caused by bonding is about 12 times larger than the exit caused by a \$10 decrease in the world oil price.<sup>35</sup>

Table 1.3 presents regression results according to firm size. Again, the results are consistent with the graphical evidence in Figure 1.6. The output quintiles are the same as in the figure. The specification in each column matches Column (3) in Table 1.2. The effect of the bond requirement is decreasing monotonically in firm size. In the first and second quintiles exit increases by about 12%, while in the third and fourth quintiles it increases by about 5%. There is no effect of the policy change on exit for firms in the largest output quintile. The background level of exit also decreases across the bottom four quintiles, from 17% in the bottom quintile to 5%. The baseline level of exit for large producers is 9.7%. This relatively high baseline level of exit in the largest quintile may be due to merger activity between large producers during this period.

### Characteristics of Exiting Firms

The model predicts that firms that shut down in response to increased insurance requirements will have dirtier environmental records due to lower safety effort. Table 1.4 describes the

---

<sup>35</sup>Of course, using marginal effects to predict the results of non-marginal changes is problematic. Recent work also suggests caution when comparing transitory price shocks with permanent policy changes (Li, Linn, and Muehlegger, forthcoming). The goal is not to precisely compare these effects, but rather to emphasize how large the effect of bonding was.

difference in environmental performance between firms that exited during the bond rollout and those that stayed.

The first column of Table 1.4 focuses on plugging of inactive wells. I do not observe well-level production, which would allow me to identify non-producing wells. However, I do observe the number and depth of wells owned and their total monthly production. For each firm, I calculate the approximate cost to safely plug all of its wells using the RRC estimate of two dollars per foot of well depth. Firms with a low ratio of revenue to expected plugging costs are more likely to have inactive wells that should have been plugged (once wells are plugged and remediated, they are no longer counted in the firm's number of wells). A low ratio of revenue to plugging costs also makes firms poor credit risks from the point of view of a bond issuer. A firm is less likely to invest in safety effort if the expected future revenue from the business is small. Particularly, it will be more attractive to exit and leave wells unplugged. There is anecdotal evidence that firms with a low ratio of revenue to expected plugging costs received high bond price offers (Texas House of Representatives, 2002). In the first column of Table 1.4, the dependent variable is an indicator variable equal to one if the firm's expected plugging costs at its most recent license renewal were more than twice its average annual revenue during 1996–2001. The sample contains firms in the industry at the end of 2001. There is one observation per firm.  $1[Exit]_i$  is equal to one for firms that exited during 2002. The constant term shows that 9.5% of firms that stayed in the industry after the bonding rule had plugging costs more than twice their average annual revenue. Among exiters, on the other hand, 22.1% of firms had these high plugging costs (9.5% + 12.6% = 22.1%).

The final two columns of Table 1.4 examine the difference in environmental rules violations for exiting vs. staying firms. I compare the number of citations during 1996–2001, normalizing by the number of wells operated during this period. Because of the count nature of the data and the large number of zeros, a count model is most appropriate for estimation. The middle column shows a negative binomial regression of rules violations on  $1[Exit]_i$ . The estimated coefficient on  $1[Exit]_i$  is positive and statistically significant, indicating that exiters were more likely to have had rules violations. The difference between exiters and stayers in the expected count of rules violations per well is given by  $e^\beta - 1$ , where  $\beta$  is the coefficient estimate for  $1[Exit]_i$ . This calculation shows that exiting firms had 42% more rules violations per well than did firms that stayed in the industry. For robustness, I show an OLS version of the same regression in the final column. The difference in expected violations per well is very similar in the OLS and negative binomial specifications.

## Oil and Gas Production

Another prediction of the theoretical model is that judgment-proof firms will reduce their output in response to increased bond requirements. I measure the effect of the bond mandate on firm-level oil and gas production using an event study design that compares firms that have already purchased a bond to firms that have not yet had to purchase a bond. Unlike exit, which is only observed once per year, oil and gas production are observed every month

for every firm. This design leverages the panel nature of the data through firm and time fixed effects, which reduce noise and allow me to better measure changes in production.

I run the following regression separately for oil and natural gas production,

$$\ln(\text{Production})_{it} = \gamma + \psi 1[\text{Bonded}] + \delta_i + \tau_t + \nu_{it} \quad (1.6)$$

The sample is limited to the 12 months during 2002, and to firms who stayed in the industry after the bond mandate.  $1[\text{Bonded}]_{it}$  is an indicator variable equal to one for firm-month observations where the license renewal date has already passed.  $\delta_i$  is a firm fixed effect, and  $\tau_t$  is a month fixed effect. In my preferred specification, I interact  $1[\text{Bonded}]_{it}$  and the month fixed effects with a categorical variable for firm size quintile. This estimates the effect of bonding separately for each size group. It also allows for separate arbitrary time trends in production within each output quintile.

The identifying assumption in this analysis is that  $\nu_{it}$  is independent of  $1[\text{Bonded}]_{it}$ , conditional on  $\delta_i$  and  $\tau_t$ . As I explained in Section 1.5,  $1[\text{Bonded}]_{it}$  is entirely determined by the firm's assigned license renewal date. So, under the assumption that license renewal dates are exogenous with respect to oil and gas production, the regression above consistently estimates the effect of bonding on output.

As a check on this assumption, Appendix Table A.1 compares the number and size of firms with license renewal dates in each month. There does not seem to be any pattern in mean output across months. In an F test of the null hypothesis of joint equality across all 12 group means, the F statistic is 0.97 (p-value 0.47).<sup>36</sup> Thus, there does not seem to be any evidence of systematic differences in firms based on the month in which they renew their operating licenses. As an additional robustness check, when I show the results of the production analysis I also show results from a placebo analysis for the year prior to the bond policy change. If differences between renewal month groups introduce bias, it should be apparent in this placebo analysis.

### Oil and Gas Production: Graphical Results

Figure 1.7 visually demonstrates the effect of the bond expansion on firm-level oil production for the smallest 80% of firms. The plotted regression coefficients and 95% confidence intervals correspond to months-from-license-renewal dummy variables. Month 0 is the firm's assigned license renewal month, which varies from January to December 2002; month -1 is the month before, and so on. The sample is limited to firms that remained in the industry after the bond requirement, and further limited to the the smallest 80% of producers, based on firm-level oil and gas output (in barrel of oil equivalents) during 1998–2001. As in Equation 1.7, firm fixed effects and month-by-revenue-quintile fixed effects are included. The figure shows no apparent trend in production prior to license renewal. Once firms renew their operating licenses, there is a clear decrease in production of about 5%. Individual

<sup>36</sup>Because the skewed nature of the data may reduce the power of the F-test, I also perform a nonparametric Kruskal Wallis test. The p-value from this test is 0.86.

event month estimates are not statistically different from one another; however, as I show below, I can reject the null hypothesis that production during the post-event months equaled production during the pre-event months for this group of firms.

### Oil and Gas Production: Estimation

Tables 1.5 and 1.6 describe the results of equation 1.7 for firm-level oil and natural gas production, respectively. Table 1.5 focuses on oil. Column (1) is a regression of logged monthly oil production on  $1[Bonded]_{it}$ , with firm fixed effects. In this specification, bonding reduces oil production by 4.7% on average across all firms. Column (2) adds month-by-output quintile fixed effects to control for variation over time in other determinants of oil production. This allows for different arbitrary trends for producers in different quintiles. The time fixed effects reduce the estimated effect to 3.2%. In Column (3), I estimate the effect of bonding separately by firm size. Output quintiles for this table are based on average annual production of oil and gas (in barrel of oil equivalents) during 1997–2001. The quintile breaks are given in the left column of the table. The largest effects are for firms in the bottom two quintiles. Bonding reduces oil production among these three groups by 8% and 7%, respectively. In the third and fourth quintiles, bonding has a small and statistically insignificant negative effect. The effect of bonding in the largest quintile is positive, although not statistically significant. A positive effect for large producers may indicate consolidation of the industry as large producers acquire the wells of exiting small firms.

Columns (4) and (5) of Table 1.5 show a placebo test using data from the previous year (2001) as a check on the identifying assumptions. If oil production were affected by some feature of license renewal other than the bond rollout, or if systematic differences between renewal month groups introduced correlation between  $1[Bonded]_{it}$  and unobserved determinants of production, we would expect to see similar results in previous years as well. There is no statistically significant effect of license renewal on oil production in the placebo analysis.

Table 1.6 addresses natural gas production. Unlike for oil production, there is no estimated effect of bonding on natural gas production in Columns (1) or (2). In Column (3), there are no statistically significant effects and there does not seem to be any pattern across quintiles in the effect of bonding. As with oil, the placebo analysis in Columns (4) and (5) shows no effect.

One potential explanation for the different results for oil and gas may be that oil is more amenable to “fly-by-night” production models. For example, gas wells must be connected to gathering pipelines in order to deliver product to market, while some oil wells can simply store oil in tanks at the lease to be picked up by truck. Consistent with that story (but not exclusively that story) is the fact that firms that produce at least some gas are larger on average than firms that produce only oil.

## Environmental Outcomes

The theoretical model predicts that increasing bond requirements will decrease environmental incidents. This section examines the effect of increased bond requirements on orphan wells, violations of water protection rules, and well blowouts. I compare outcomes before and after the 2001 policy change. The visual evidence suggests that there were decreases that coincided in time with the increased bond requirement. In the econometric analysis, I run regressions of the form,

$$Y_{it} = \zeta + \phi 1[After]_t + Month_t + u_{it} \quad (1.7)$$

$Y_{it}$  is an environmental outcome; for example, the monthly count of rules violations at the firm level.  $1[After]_t$  is an indicator variable equal to one after the policy change. To allow for anticipatory increases in safety effort by firms trying to lower their expected bond premiums,  $1[After]_t$  becomes equal to one in June 2001, when the bill requiring the higher bond amounts passed the Texas Legislature.<sup>37</sup>  $Month_t$  is a parametric time trend, and  $u_{it}$  is an error term capturing unobserved determinants of the outcome.

Because environmental incidents are rare, several years of outcome data are needed to make empirically meaningful econometric comparisons. Relative to the analysis of producer exit, which examined a small neighborhood around the implementation threshold, and of output, which limited the sample to 2002 and compared already-bonded to not-yet-bonded firms, the analysis of environmental outcomes requires stronger identifying assumptions. It is not possible to definitively rule out other shocks that could have driven some of the observed changes. However, I show that there do not seem to be other regulatory changes that would explain these large changes in environmental outcomes. In addition, for one outcome where similar data are available, I show that outcomes do not change similarly in a neighboring state.

### Environmental Outcomes: Graphical Results

Figure 1.8 shows the share of operators leaving the industry each quarter that left behind orphan wells. Orphan wells are unplugged wells for which no registered operator exists. Firms that exited prior to the end of the twelve-month rollout period would not have been bonded, while firms exiting after the implementation period would have been. There is also little reason to expect anticipatory safety effort in response to the policy change, since orphaning occurs among firms exiting the industry. The vertical line represents December 2002, the end of the rollout period. The data are noisy, but there is an apparent decrease in the share of operators leaving orphan wells at the same time that the increased bonding requirement takes effect. The light gray horizontal lines represent the mean rate of orphaning during five years before and five years after the policy change (in the estimation, I also allow for time trends). The mean rate of orphaning after the policy change is lower by about 50%.

---

<sup>37</sup>Anticipatory safety effort may also affect firm-level oil and gas production. If so, this would lead me to underestimate the effect of bonding on output in Section 1.5.

Because firms may orphan different numbers of wells, Appendix Figure A.3 shows the total number of wells orphaned each quarter before and after policy change. The total number of wells orphaned falls substantially.

Figure 1.9 shows the total number of water protection rules violations each quarter. These are violations of Statewide Rules 8 and 14, which govern water protection as described in Section 1.4. The vertical dashed line shows the beginning of the policy implementation period in January, 2002. The lighter dotted line shows the bill’s passage in June, 2001. Again, the horizontal gray lines show the mean number of violations during five years before and five years after the policy change. The data are noisy. The number of violations after the policy change is lower, but it is difficult in the figure to separate discrete changes from what may be a smooth downward trend across the implementation threshold.

Figure 1.10 shows well blowouts. Because blowouts occur most frequently during drilling, in this figure I normalize the number of blowouts by the number of active drilling rigs each quarter.<sup>38</sup> As in the previous figure, the vertical lines represent the policy passage and implementation, and the horizontal lines represent five-year averages before and after the policy change. Before 2001, the time series of blowouts is noisy but relatively flat. There are no clear discrete changes. In 2001, there is a sharp drop in blowouts coincident with the passage of the universal bond mandate. The blowout rate stays low following this change through the end of the panel.

The time series evidence in these figures is suggestive, but it does not rule out other shocks to the oil and gas industry that could have caused these changes. It is not possible to completely overcome this limitation of this portion of the analysis, but it is possible to look for obvious alternative explanations. As one simple robustness check, Figure 1.11 shows the time series of onshore blowout rates in Louisiana, a nearby state that is also a large oil and gas producer. If blowout rates in Louisiana changed similarly to Texas blowout rates in 2001, that would suggest that the change was related to national-scale shocks to the oil and gas industry, as opposed to Texas regulations. The Louisiana blowout data are from the Louisiana Department of Natural Resources, and, as in the Texas figure, the number of active drilling rigs each month comes from the Baker Hughes Historical Rig Count dataset.<sup>39</sup> The vertical line shows June, 2001, and the horizontal gray lines show five-year mean blowout rates before and after. Perhaps due to lower production, the data prior to 2001 are noisier. There are more quarters with many blowouts, but also more quarters with no blowouts. There may also be a smooth downward trend over time, although the mean blowout rate during the post-period is relatively similar to the pre-period. However, there is not a visually obvious discrete change in the blowout rate in 2001 like there is for Texas.

<sup>38</sup>Data on the number of active rigs each month come from the Baker Hughes Historical Rig Count dataset, <http://www.bakerhughes.com/rig-count>.

<sup>39</sup>Louisiana blowout data are from the Department of Natural Resources SONRIS database: <http://www.sonris.com/dataaccess.asp>. To focus on onshore production, I include all blowouts at wells with a valid Public Lands Survey System “Township” designation. The number of active drilling rigs each month comes from the Baker Hughes Historical Rig Count dataset (<http://www.bakerhughes.com/rig-count>). I include rig activity for “land” and “inland water” areas.

To consider whether improvements in environmental outcomes may have been caused by other Texas regulatory changes, Appendix Table A.3 shows all of the rules that were implemented or amended by the RRC in 2001 and 2002. I include these years to allow for immediate responses to implementation of new regulations, as well as anticipatory responses to regulations announced in 2001 but implemented in later years. The first column shows the date that the proposed action was published for public comment in the Texas Register. The second column shows the date that the regulation took effect. The other RRC actions were primarily procedural, and it seems unlikely that any of them would have driven the observed changes in environmental outcomes. In summary, there do not seem to be obvious confounders that would explain the changes in environmental outcomes that I observe.

### Environmental Outcomes: Estimation

Table 1.7 shows regression estimates of the change in environmental outcomes coincident with the policy change. The first two columns show a linear probability model focused on orphan wells. The sample in these two columns includes one observation per firm, in its month of exit. The dependent variable is an indicator variable equal to one if any of the firm's wells appear on the March 2014 version of the RRC Orphan Well List. 1[*After*] is an indicator variable equal to one in June 2001 and all later months. In Column (1), a simple before/after comparison estimates a 4.4 percentage point decrease in the probability that firms orphan wells upon leaving the industry. This is a 56% decrease from the pre-bonding rate of well orphaning of 7.8%. Column (2) adds a quadratic time trend and month-of-year fixed effects. This reduces the estimated effect slightly to 3.2 percentage points, or a 51% decrease. This effect is consistent with the visual evidence in Figure 1.8.

Columns (3) and (4) address environmental rules violations. Because of the count nature of the data, I use negative binomial regression.<sup>40</sup> The dependent variable is the count of rules violations each month. This is normalized by the number of wells operated by the firm that month. Column (3) shows that a before/after comparison estimates a 14% decrease in the frequency of rules violations per well after the policy change. In Column (4), including an exponential time trend and month-of-year fixed effects increases the estimate slightly, but decreases its statistical significance. Again, the estimated effects are consistent with the graphical evidence in Fig 1.9.

Columns (5) and (6) address well blowouts. Again, I use negative binomial regression because of the count nature of the data. In these regressions the count of well blowouts in each month is normalized by the number of wells being drilled by the firm in that month. There is a 44% decrease in blowouts in the before/after comparison, and a 67% decrease after controlling for an exponential time trend and month-of-year fixed effects.

The theoretical model predicts that changes in environmental outcomes will result from changes in industry composition (extensive margin changes) and increases in cost internalization by firms that remain in the industry (intensive margin changes). It is interesting to

---

<sup>40</sup>For both rules violations and well blowouts, results are similar for poisson and negative binomial models. The negative binomial is preferred because it accommodates overdispersion.



compare the magnitudes of these effects. Table 1.8 measures the intensive margin effect for rules violations and well blowouts (orphan wells, by definition, apply only to exiting firms). This table repeats the regressions from Columns (3)–(6) of Table 1.7, limiting the sample to firms that entered the industry before December 2001 and exited after January, 2003 (*i.e.*, firms that stayed in the industry through the bond implementation period).

I focus on the specifications that include time trends and month-of-year fixed effects. For rules violations, Column (2) shows that the intensive margin effect was about a 16% decrease in the average rate of rules violations per well. This is less than the 19% overall decrease when exiting firms are included, which means that exiters had more accidents during the pre-period (This is consistent with the findings in Table 1.4). But the intensive margin effect is still large. The results for blowouts are similar. Column (4) shows that the intensive margin effect is a 62% decrease, which is smaller than the overall decrease of 67%, but still large.

## 1.6 Discussion

The bond requirement reduced the number of small firms and their output, with no effect on larger firms. At the same time, there were substantial improvements in environmental outcomes. These results are consistent with greater cost internalization by judgment-proof firms in response to the bond requirement. Prior to the policy change, financially weak operators could produce oil and gas at a low private cost by exerting minimal safety effort and avoiding any environmental costs through bankruptcy. Bonding made this business model less attractive. Very weak firms had little incentive to exercise safety effort, since it was clearly not in their private interest. Thus, they received high price offers from insurers that effectively pushed them out of the industry. At the same time, operators who became bonded were pressured by insurers to operate safely. Investing in safety reduced the private cost advantage of small firms, so that many sold their assets and left the industry.<sup>41</sup> These changes in the composition of firms and the privately optimal level of safety effort led to decreases in orphan wells, rules violations, and well blowouts.

### Alternative Explanations

#### Credit Market Inefficiencies

If the surety bond market did not operate efficiently, firms may have received high bond price offers for reasons unrelated to environmental risk. Given the environmental improvements observed, it seems unlikely that bond prices and environmental risk were completely unrelated. Nevertheless, it is worth considering to what degree credit market inefficiencies also contributed to exit. Many insurers and banks offered bonds, so it is unlikely that in-

---

<sup>41</sup>Of the leases belonging to firms that exited in 2002, about 88% were transferred to other operators at the time of exit.

urer market power raised bond prices (Gerard and Wilson, 2009). Underwriting expenses do introduce transaction costs, but the qualitative evidence does not suggest that underwriting costs were a primary driver of producer exit. As discussed, newspaper accounts and legislative records describe a small share of firms receiving very high bond price offers, while typical price offers were only 1-3% of the face value of the bond.

The final potential source of credit market inefficiency is asymmetric information about accident and default risk. Classic models of asymmetric information in credit and insurance markets raise the possibility of excess demand in equilibrium (Jaffee and Russell, 1976; Rothschild and Stiglitz, 1976). The intuition behind these models is that the inability to distinguish high-risk and low-risk types leads the bank or insurer to not offer contracts that would be profitable with the low-risk type, but would be particularly attractive to high-risk types. More recent empirical work has found that supply-side responses to asymmetric information can allow lenders and insurers to serve risky client groups. Adams, Einav, and Levin (2009) and Einav, Jenkins, and Levin (2013) document how sophisticated credit scoring systems and down payment requirements allow lenders to operate in the market for subprime car loans. Analogous strategies in oil and gas surety bond markets that likely mitigated credit market inefficiencies include credit scoring, availability of detailed information on firms' compliance histories from the RRC, and collateral requirements.

### **Inefficiently High Bond Requirements**

As discussed in Section 1.3, when bond requirements are high, transaction costs created by collateral requirements can prevent participation by responsible firms. In this case, however, bonds were set far below maximum possible damages. The required bond face values were several thousand dollars per well, or less. While it is difficult to value the damages from a worst-case groundwater or surface water contamination event, they are almost certainly much larger than that.<sup>42</sup> And the costs from a bad well blowout can reach tens of millions of dollars (Jones, 2003). With bond levels so far below maximum possible damages, it is unlikely that collateral costs kept out firms that otherwise would have operated safely.

### **Welfare**

The model in Section 1.3 provides a framework for evaluating the welfare effects of the expanded bond requirement. The precise welfare impacts depend on several highly uncertain parameters, such as the cost of a groundwater contamination incident. Nevertheless, it is relatively clear that the sign of the welfare effect is positive. This policy change caused

---

<sup>42</sup>Some studies have reported cleanup costs for groundwater contamination incidents, but these values are highly uncertain and highly variable. National Research Council (2013) finds that the average cost to remediate a contaminated groundwater site under the US Resource Conservation and Recovery Act is \$11.4 million dollars, and that the average cost to remediate a leaking underground storage tank is \$125,000. Greenstone and Gallagher (2008) report that the average cost to remediate a "Superfund" site is \$43 million (but that the average value of cleanups is lower).

firms to internalize a negative externality to a greater extent than they had previously. This reduced the frequency of costly environmental damages. It also appears to have re-allocated production from small judgment-proof producers to potentially more efficient producers, since 88% of the leases belonging to exiting firms were transferred to other producers.

The primary costs created by the policy were increased private costs of safety effort for oil and gas producers. If fines for environmental incidents are set at or below true social damages, profit maximization implies that the net difference of environmental benefits and effort costs is positive. If the marginal cost of safety effort exceeded the expected benefit in terms of reduced environmental costs, firms would not invest further in safety. In practice, expected fines are likely lower than the true social costs because of difficulties in detection, attribution, and enforcement discussed in Section 1.3 and Shavell (2007).

There are additional costs due to underwriting expenses and collateral requirements. These should be subtracted from the benefits of increased safety to calculate the true net welfare change. However, they are unlikely to outweigh the environmental benefits. Collateral requirements for firms perceived to be safe producers are small.

In another industry, the welfare calculation might be affected by competitiveness concerns or by the fact that some judgment-proof producers had lower true production costs. In oil and gas, neither of these is likely to play a large role. Crude oil and natural gas are commodities, and even after the bond mandate, producers in Texas faced essentially perfect competition both within the state and from the world market. There is also no reason to think that exiting firms were more efficient than producers that remained; instead, as argued previously, they were likely to have been higher-cost firms.

In fact, there is an argument to be made that bond requirements in Texas are still below the optimal level. Calculating the optimal bond amount requires estimates of the value of damages,  $h$ , the effort elasticity of damages,  $\gamma'(x)$ , and the capital costs imposed on firms by tying up collateral,  $r$ . These are highly uncertain. However, the empirical analysis showed that the elasticity of environmental damages over the observed change in bond requirements was large. The effect on total production was small, as output was re-allocated to other (potentially more efficient) producers. If further modest increases in bonds were to yield similar environmental improvements, the benefits would likely be larger than the costs.

## 1.7 Conclusion

This paper focuses on a historical case study, but the results are relevant today. Between 2006 and 2013, U.S. oil production increased by 65% and natural gas production by 40%, largely due to hydraulic fracturing. In 2015, U.S. crude oil production is expected to be at its highest level since 1970.<sup>43</sup> The oil and gas boom has had economy-wide benefits. At the same time, it presents environmental challenges on a massive scale. The large number of

---

<sup>43</sup>Oil and gas production growth are calculated from U.S. Energy Information Administration (EIA) “U.S. Field Production of Crude Oil” and “U.S. Natural Gas Gross Withdrawals” data. The 2015 forecast of 9.5 million barrels per day is from the EIA Short-Term Energy Outlook, October 2014.

projects being developed creates more risk of accidents, and the deployment of new chemicals and techniques creates novel risks.

The U.S. oil and gas industry still includes many small firms. Davis (2015) demonstrates the lack of concentration in onshore natural gas drilling. A back-of-the-envelope calculation suggests that the majority of firms drilling new gas wells in 2012 had annual revenues of several million dollars or less. This means that their effective liability exposure is much less than the damages from a major groundwater contamination incident, well blowout, or spill into surface water.

At the same time, bond requirements in most jurisdictions remain very low. The minimum bond requirements for oil and gas production on federal lands have not been increased since 1960, even to adjust for inflation. The Texas requirements examined in this paper are some of the highest among major oil- and gas-producing states (Appendix Figure A.4), and they are still small compared to maximum potential damages. The results of this study support arguments to increase bonds in other jurisdictions at least to the amounts required in Texas. While it is impossible to extrapolate beyond the observed bond levels, it also seems likely that somewhat higher bond requirements would yield further benefits given that Texas' requirements are still well below potential damages.

More broadly, this paper extends our understanding of bankruptcy and market structure in dangerous industries. The theoretical model extends existing models to formalize the relationship between the judgment-proof problem and industry structure. The empirical analysis validates the theoretical predictions. The judgment-proof problem inflated the number of small producers, and led them to produce more than was efficient while exerting less-than-efficient safety effort. The analysis departs from existing cross-sectional studies through its credible quasi-experimental research design and its use of detailed firm-level administrative data.

The results suggest that bankruptcy should be taken seriously as a determinant of market structure in industries with substantial liability exposure. Within the energy sector, this has implications for transportation of oil, natural gas, and gasoline and other refined products by pipeline, road, and rail.<sup>44</sup> Other examples of sectors with small firms and high liability risk include chemical manufacturing, transportation network companies, and retail gasoline. More work in other settings will help to gauge the generality of these results, but the findings of this study demonstrate the important incentive effects of bankruptcy in dangerous industries. Continuing to evaluate and address this market failure will be an important component of efficient safety regulation in some of the world's most important industries.

---

<sup>44</sup>In July 2013, a train carrying crude oil derailed and exploded in Lac-Mégantic, Quebec, killing 47 people. The total damages were estimated at \$200 million. The railroad, which had a \$25 million liability insurance policy, declared bankruptcy. Morris, Betsy. "Fiery Oil-Train Accidents Raise Railroad Insurance Worries." *Wall Street Journal*. January 8, 2014.

Figure 1.1: Long-Run Average Cost Functions In a Hazardous Industry

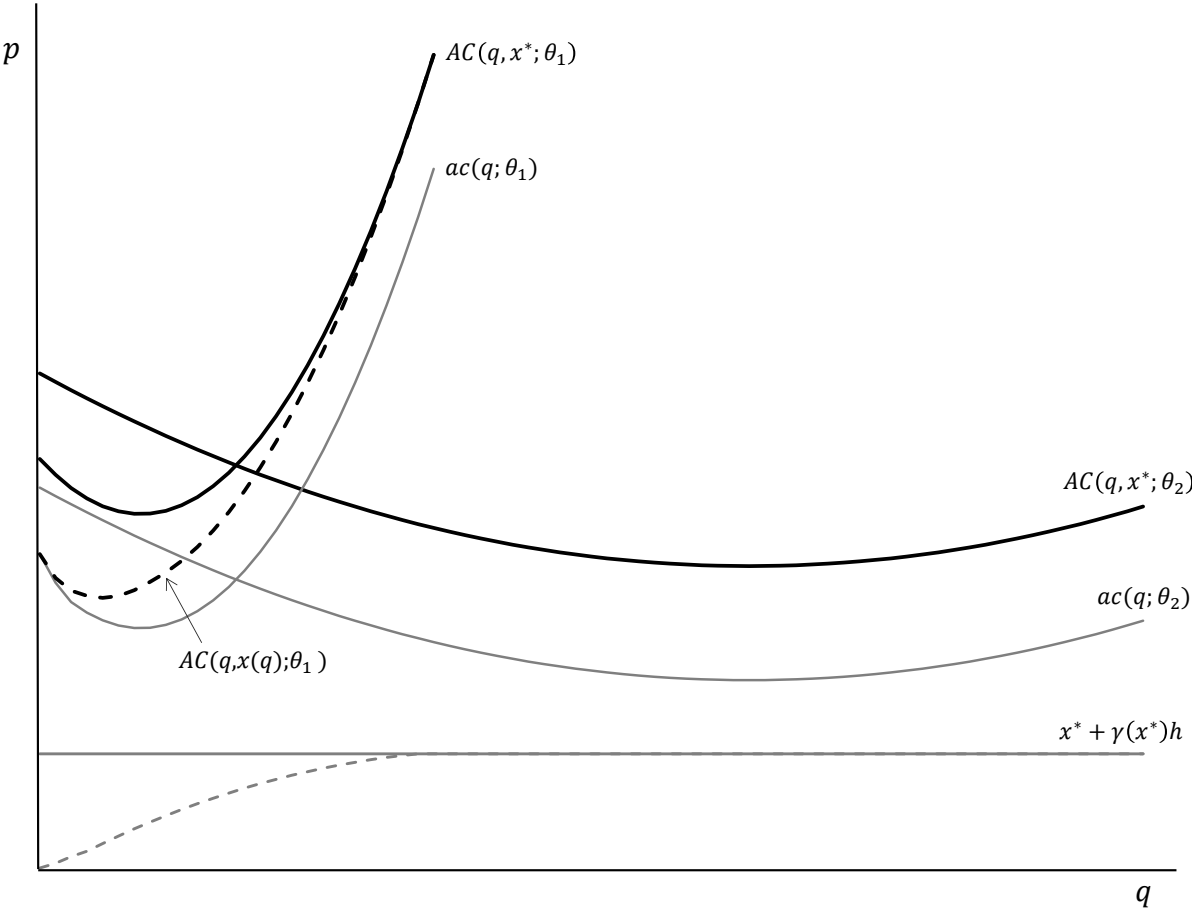
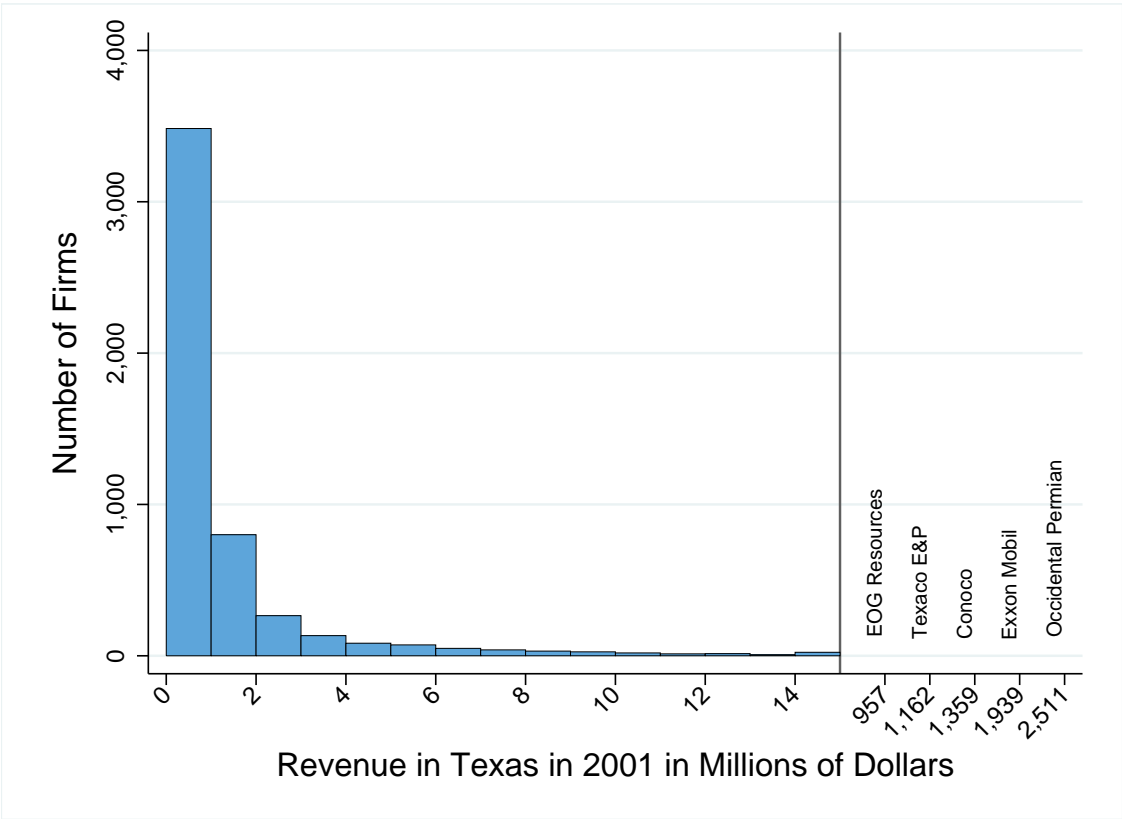
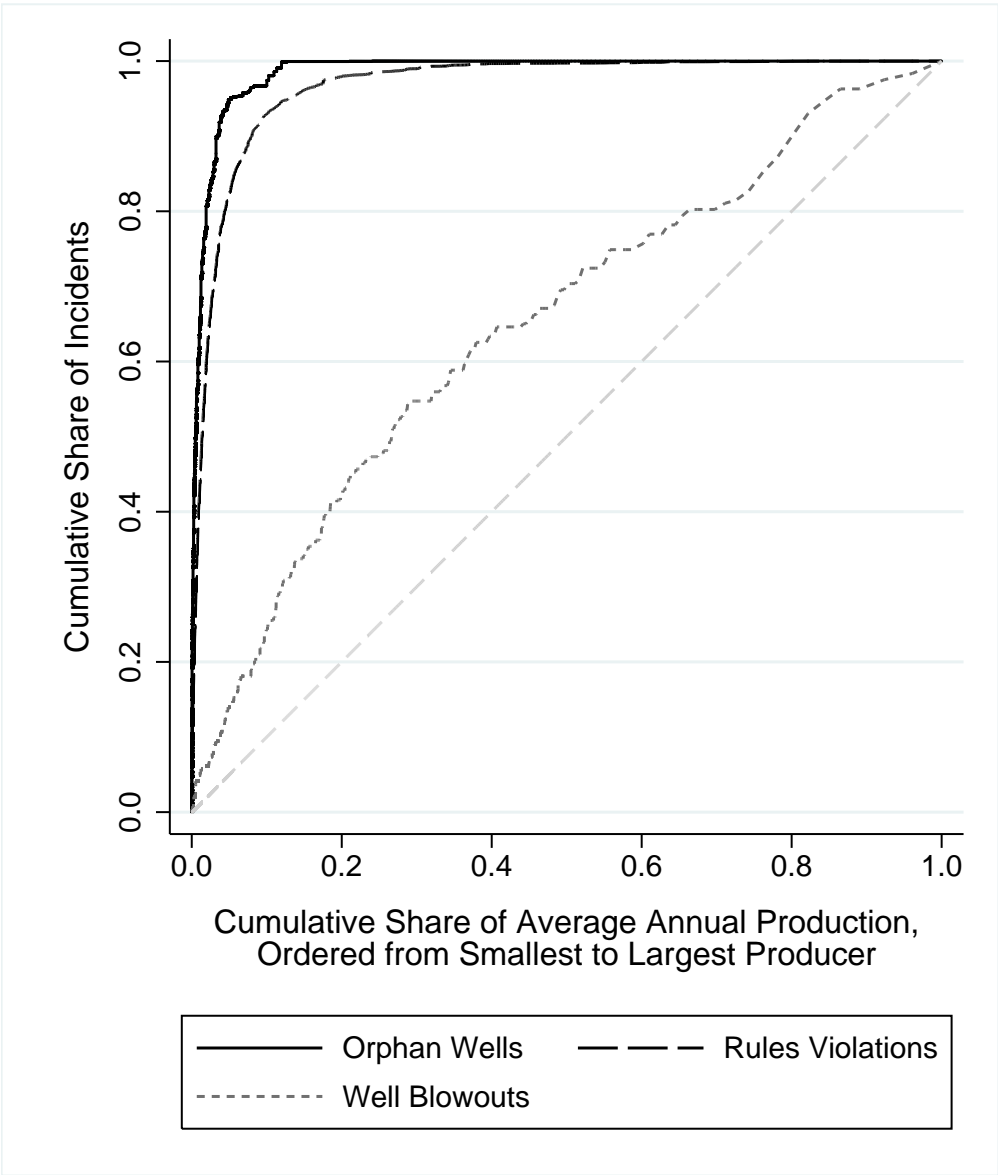


Figure 1.2: Size Distribution of Oil and Gas Producers in Texas



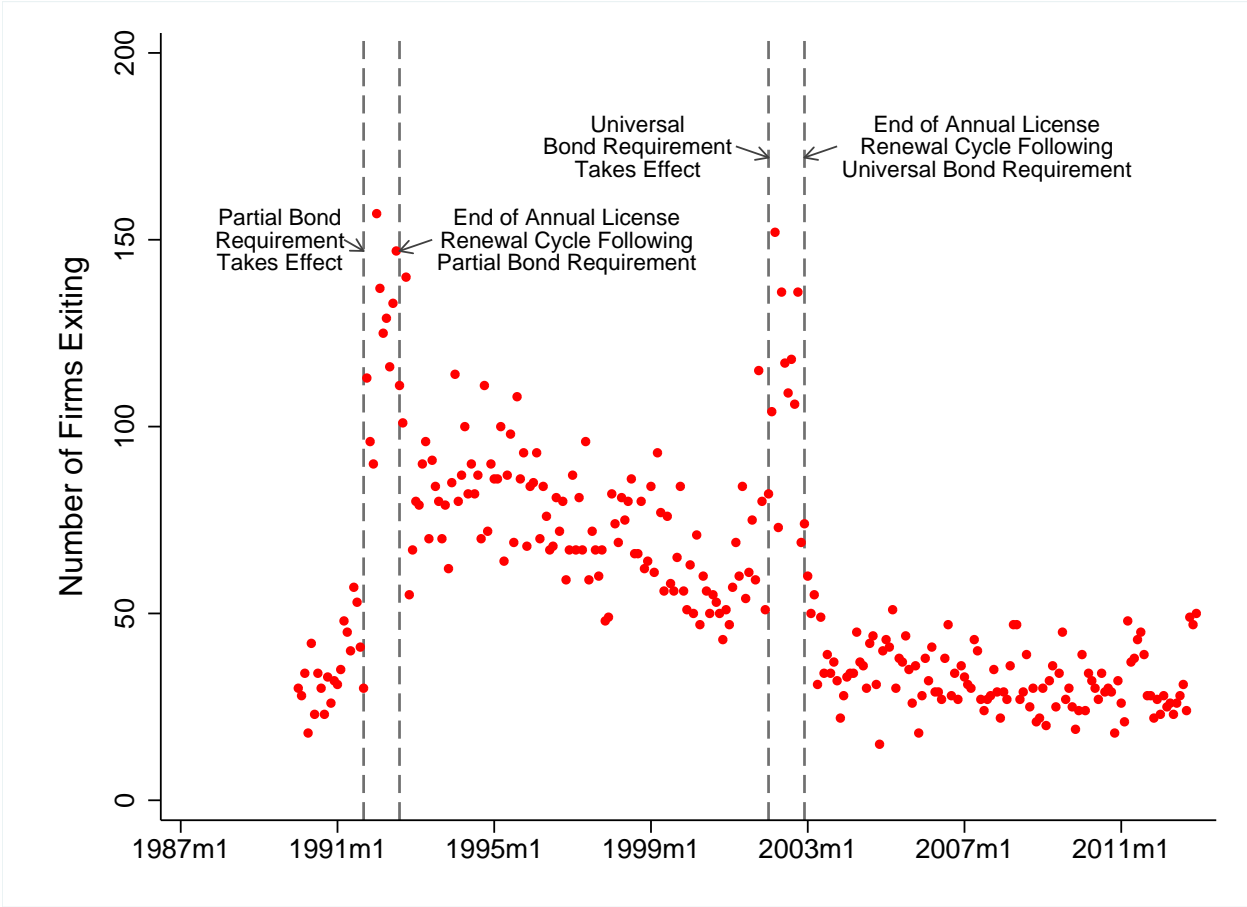
Notes: This figure shows the distribution of production revenues in 2001 for the 5,302 oil and gas producers that were in the industry for the full year. Revenues are calculated using EIA Texas first purchase prices for oil and EIA Texas wellhead prices for natural gas. Dollar amounts are in 2010 dollars. The horizontal axis between 15 million and 957 million dollars is not shown; there are 239 firms in this range.

Figure 1.3: Cross-Sectional Comparison of Environmental Incidents and Firm Size



Notes: This figure shows that environmental incidents are concentrated among the smallest firms. The figure uses data from 1990 to 2001. For each firm, I calculate average annual oil and gas production in barrel of oil equivalents (1 BOE = 1 barrel of oil or 6,000 cubic feet of natural gas). The horizontal axis represents the sum of average annual production for all firms. Producers are ordered from left to right on this axis, so that, for example, 0.2 represents the 20% of total annual production that comes from the smallest firms. The vertical axis represents the total number of environmental incidents during this period. So, 0.4 on the vertical axis corresponds to 40% of total incidents during 1990–2001. The gray dashed line has a slope of 1.

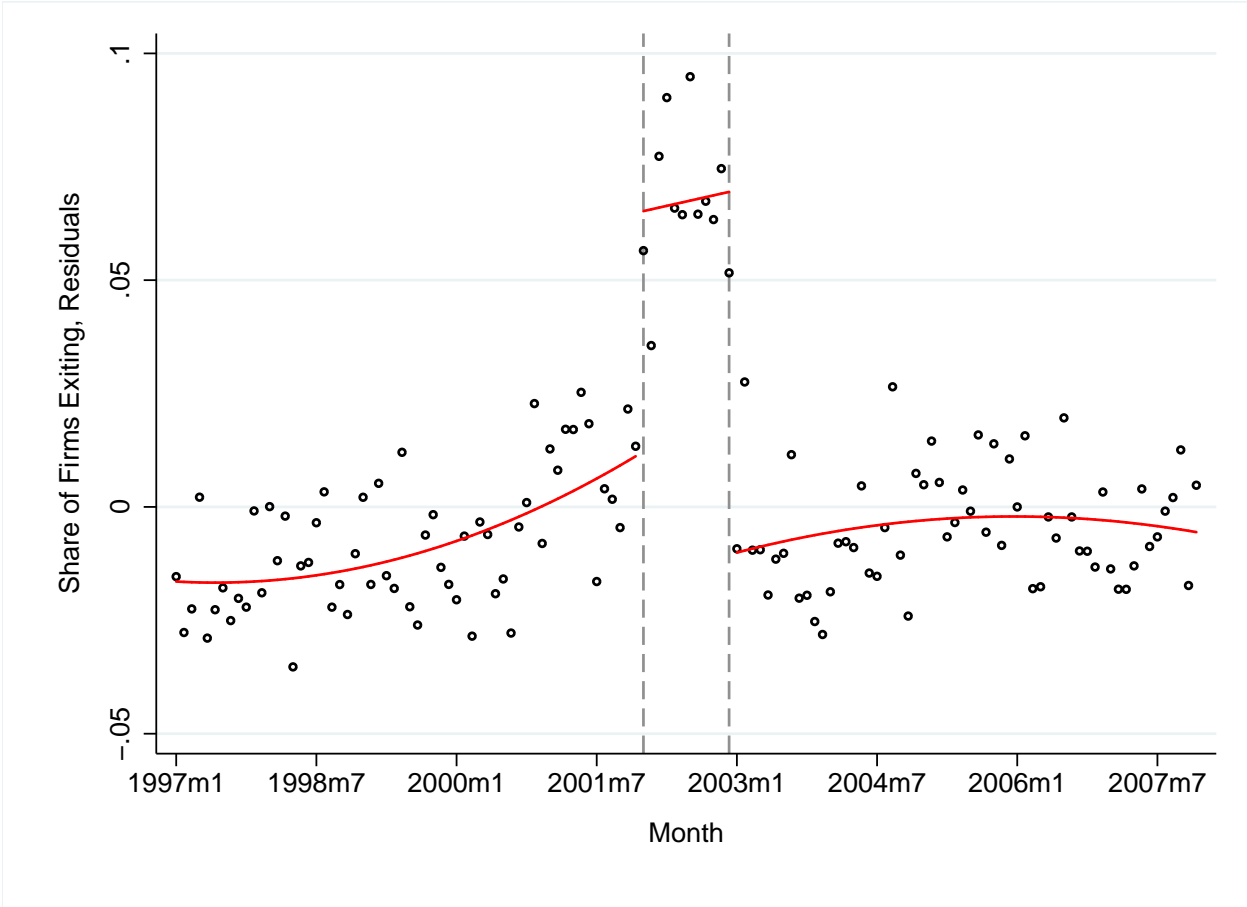
Figure 1.4: Number of Firms Exiting by Month



Notes: This figure shows the number of firms leaving the industry each month. Exit date is defined as 365 days after the firm’s final annual operating license renewal. The sample includes all firms with oil or gas leases from 1990 to 2010. The initial rollout of the partial bond requirement occurred during September, 1991 to August, 1992. The initial rollout of the universal bond requirement occurred during January–December, 2002.

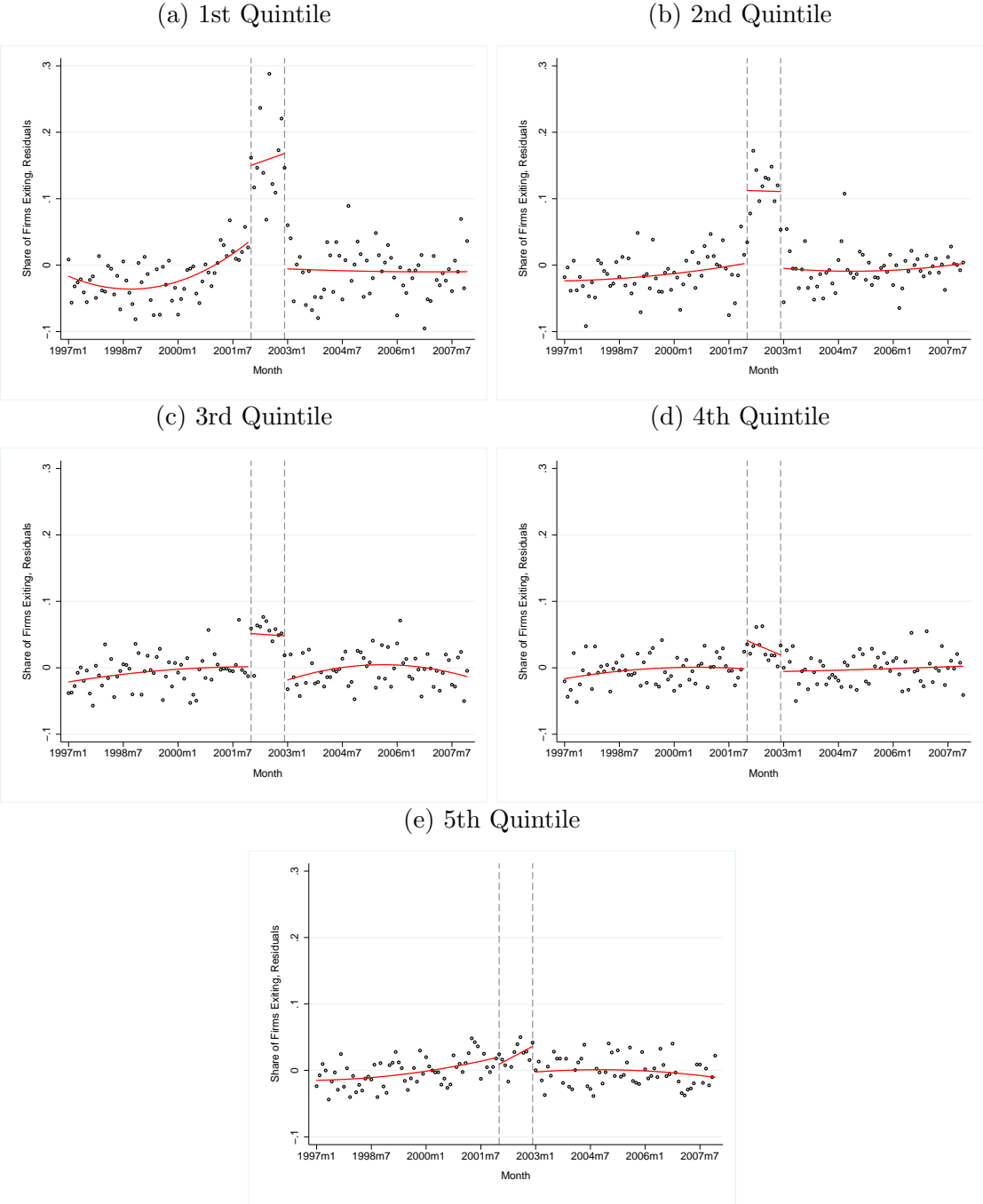


Figure 1.5: The Effect of the 2001 Bond Requirement on Exit



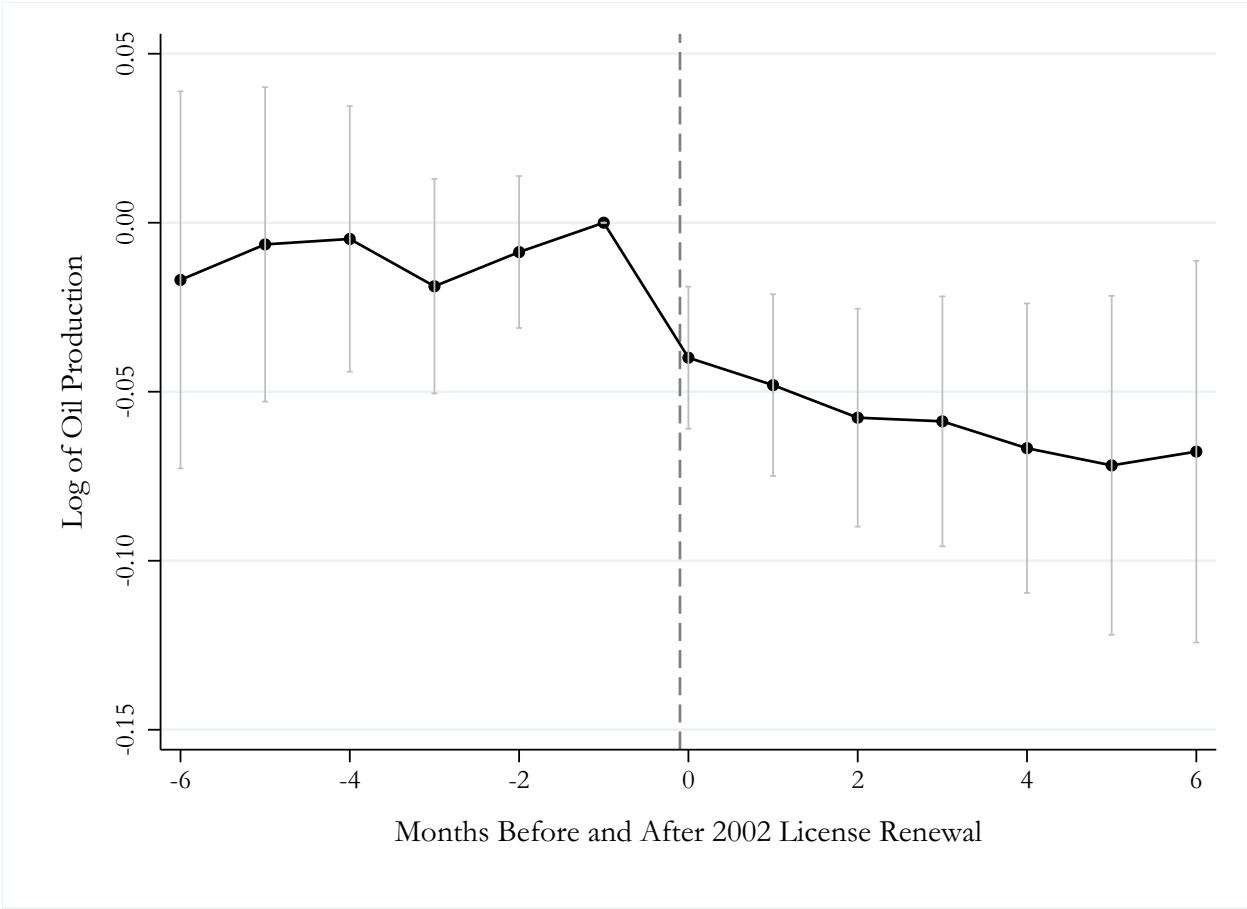
Notes: This figure shows the effect of the 2001 bond requirement on exit. The sample includes all firms that produced oil or natural gas during 1997–2006, and that reported at least six months of production during 1996–2007. There is one observation per firm per year, in the firm’s assigned license renewal month. The dots are monthly means of the residuals from a regression of an indicator variable for exit on month-of-year fixed effects and monthly crude oil prices. The red curves represent a quadratic polynomial fit for 1997–2001, a separate linear polynomial for 2002, and a separate quadratic polynomial fit for 2003–2006.

Figure 1.6: The Effect of the 2001 Bond Requirement on Exit, by Firm Size



Notes: This figure shows the effect of the bond requirement on exit by quintile of firm size. The sample and analysis are the same as in Figure 1.5. Each panel shows results from a single quintile of the output distribution. Output is calculated as average annual production for 1997 to 2006 in barrels of oil equivalent (BOE). One BOE equals one barrel of crude oil or six MCF of natural gas. Average annual production is 12 times the average monthly production across all non-zero months.

Figure 1.7: Oil Production By Small Firms Before and After Bonding



Notes: This figure shows the effect of bonding on monthly oil production for the smallest 80% of firms. The horizontal axis shows time relative to the firm’s assigned license renewal month (which ranges from January to December 2002). The sample is limited to firms that remained in the industry after the bond requirement, and to the bottom four quintiles of the output distribution. See text for details. The plotted points represent regression coefficients on indicator variables equal to one when the firm is a corresponding number of months away from license renewal. Vertical bars represent 95% confidence intervals. The omitted category is  $t=-1$ , the month before renewal. These regressions include firm and month-of-sample-by-revenue-quintile fixed effects. Revenue quintiles are defined as in Table 1.5. Standard errors are clustered at the firm level.

Figure 1.8: Share of Exiters Orphaning Wells, by Quarter of Exit

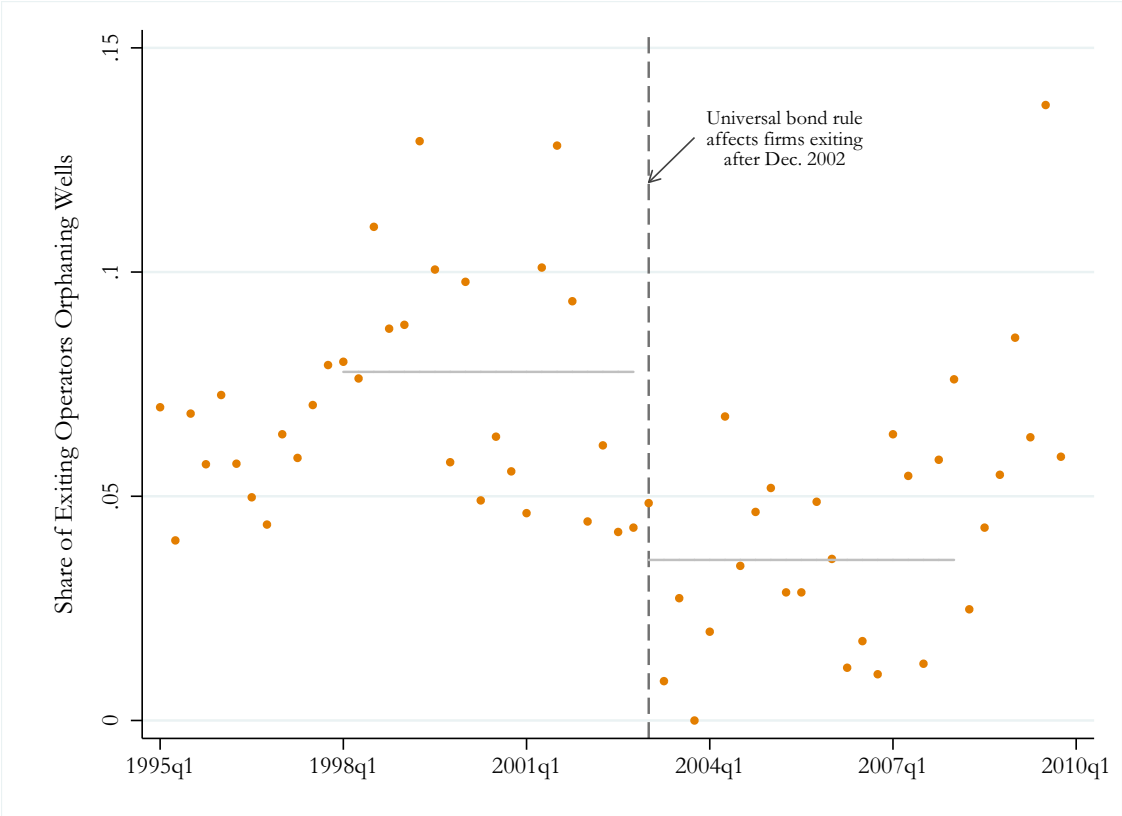


Figure 1.9: Water Protection Rules Violations by Quarter

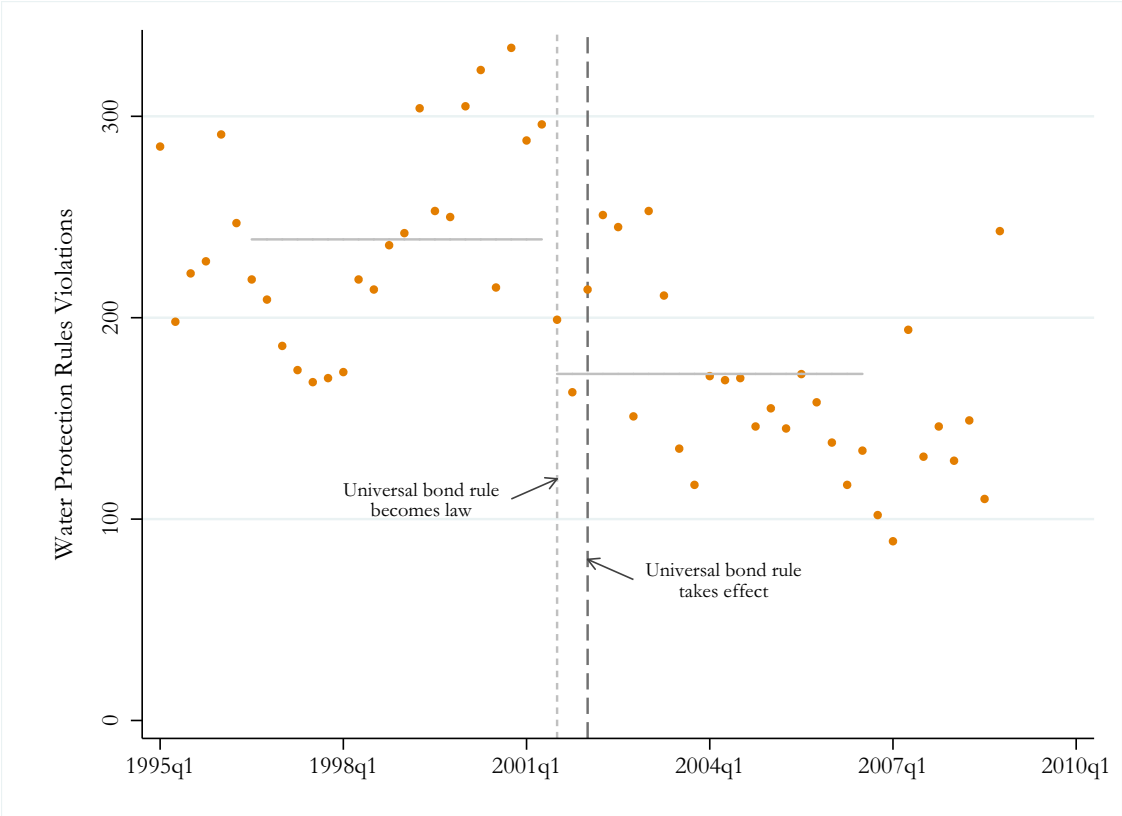


Figure 1.10: Well Blowouts by Quarter

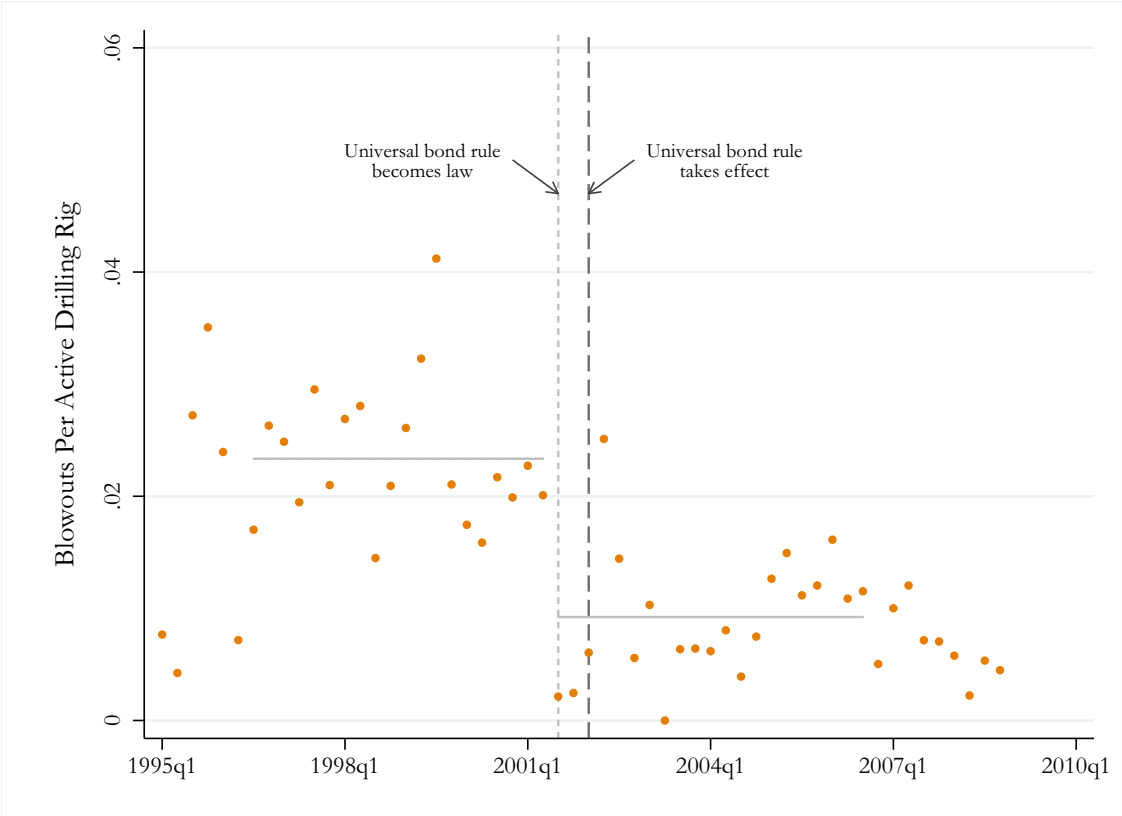


Figure 1.11: Well Blowouts in Louisiana by Quarter

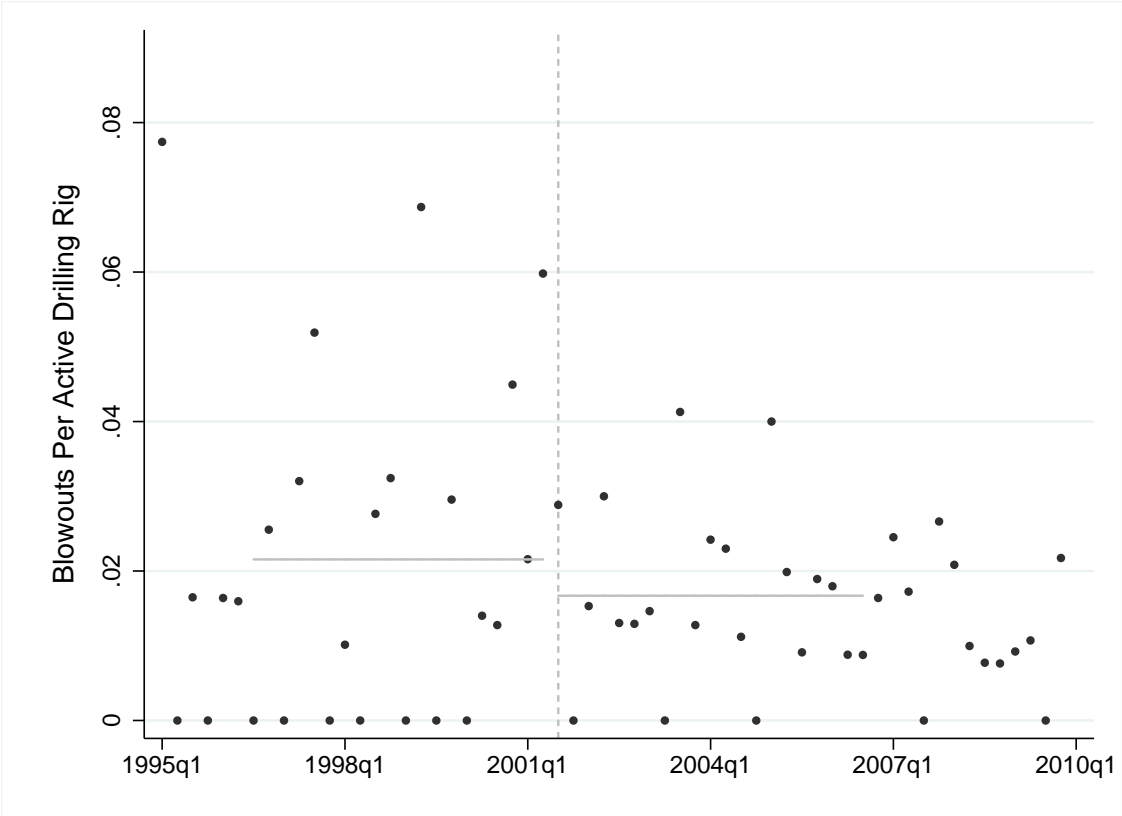


Table 1.1: Quintiles of Annual Production Revenue Per Firm in 2001

---

---

5	\$2,510,918,000
4	\$1,391,000
3	\$344,000
2	\$114,000
1	\$33,0000

---

---

Notes: This table lists quintiles of production revenue in 2001 for the 5,302 oil and gas producers that were in the industry for the full year. Revenues are calculated using EIA Texas first purchase prices for oil and EIA Texas wellhead prices for natural gas. Dollar amounts are in 2010 dollars.



Table 1.2: Effect of the Bond Requirement on Exit

	(1)	(2)	(3)
Begin Rollout	0.068*** (0.010)	0.063*** (0.010)	0.060*** (0.010)
End Rollout	-0.082*** (0.011)	-0.084*** (0.013)	-0.083*** (0.013)
Oil Price (\$100/bbl)		-0.044** (0.020)	-0.051** (0.021)
Constant	0.090*** (0.006)	0.089*** (0.005)	0.093*** (0.006)
Month-of-year FE	No	No	Yes
Local Time Polynomial	Yes	Yes	Yes
N	71,826	71,826	71,826
Firms	10,978	10,978	10,978

Notes: This table reports the results of 3 separate OLS regressions. The sample includes 1997–2007 and includes all firms with at least six months of oil or natural gas production during 1996–2007. The data are a monthly panel with one observation per firm per year, in the firm’s assigned license renewal month. The dependent variable is an indicator variable for exit in a given year. Begin Rollout is an indicator variable equal to one starting in January 2002. End Rollout is an indicator variable equal to one starting in January 2003. All specifications include a quadratic polynomial in time for 1997–2001; a separate linear polynomial for 2002; and a separate quadratic polynomial for 2003–2007. Oil prices are monthly average Texas first purchase prices in the month of license renewal, in hundreds of 2010 dollars. Standard errors are clustered by quarter. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; \* at the 10% level.

Table 1.3: Effect of the Bond Requirement on Exit by Output Quintile

	(1)	(2)	(3)	(4)	(5)
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Begin Rollout	0.121*** (0.014)	0.120*** (0.025)	0.056*** (0.017)	0.045*** (0.013)	-0.011 (0.013)
End Rollout	-0.158*** (0.029)	-0.114*** (0.022)	-0.095*** (0.018)	-0.049** (0.019)	-0.014 (0.015)
Oil Price (\$100/bbl)	-0.115** (0.048)	-0.083** (0.033)	-0.081** (0.033)	-0.030 (0.023)	0.025 (0.033)
Constant	0.170*** (0.013)	0.094*** (0.020)	0.079*** (0.012)	0.050*** (0.008)	0.097*** (0.010)
N	11,903	13,267	14,667	16,197	15,792
Firms	2,129	2,185	2,202	2,226	2,236

Notes: This table reports estimates separately by output quintile for the regression specification shown in Column 3 of Table 1.2. Each column shows results from a regression that includes one quintile of the output distribution. Output is calculated as average annual production during 1996–2007 in barrels of oil equivalent (BOE). One BOE equals one barrel of crude oil or six MCF of natural gas. Average annual production is 12 times the average monthly production across all non-zero months. All regressions include month-of-year fixed effects and the same local polynomial in time as in Table 1.2. Standard errors are clustered by quarter. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; \* at the 10% level

Table 1.4: The Relative Environmental Performance of Exiting Firms

	Plugging Costs > 2 Years Revenue	Water Protection Rules Violations	
	OLS	Neg. Binomial	OLS
1[ <i>Exit</i> ]	0.126*** (0.015)	0.348** (0.143)	6.999 (4.319)
% Difference in Expected Count	133%	42%	38%
Constant	0.095*** (0.005)	2.775*** (0.067)	18.329*** (1.415)
Normalization	None	Thousand Wells	Thousand Wells
Firms	4,868	4,868	4,868

Notes: This table reports the results of three separate regressions. The sample includes firms active at the end of 2001 with at least 6 months of production during 1996–2001. There is one observation per firm. 1[*Exit*] is an indicator variable equal to one for firms that exited during the bond rollout in 2002. In the first column, the dependent variable is an indicator variable equal to one if the estimated cost to plug all of the firm’s wells at its most recent license renewal was more than twice the firm’s average annual production revenue during 1996–2001. In the second and third columns, the dependent variable is the count of violations of Statewide Rules 8 or 14 during June 1996–June 2001. This is normalized by the average number of wells operated during the same period (in thousands of wells). For OLS, the normalization is implemented by dividing by the number of wells; for the negative binomial, the log of wells is included as a regressor with coefficient constrained to one. For OLS specifications, the % difference between groups is the regression coefficient  $\beta$  divided by the mean outcome for the non-exiting firms. For the negative binomial, the % difference is  $\exp(\beta) - 1$ . See text for details. Standard errors are Eicker-White heteroscedasticity-consistent. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

Table 1.5: Effect of Bonding on Oil Production for Remaining Firms

	Implementation Year			Placebo Year	
	(1)	(2)	(3)	(4)	(5)
1[Bonded]	-0.047*** (0.009)	-0.032*** (0.010)		0.001 (0.011)	
1[Bonded]*Quintile 1 (1,800 BOE)			-0.079*** (0.030)		0.006 (0.034)
1[Bonded]*Quintile 2 (5,700 BOE)			-0.074*** (0.025)		-0.024 (0.023)
1[Bonded]*Quintile 3 (16,100 BOE)			-0.011 (0.022)		0.028 (0.022)
1[Bonded]*Quintile 4 (60,200 BOE)			-0.027 (0.020)		-0.009 (0.018)
1[Bonded]*Quintile 5 (60,000,000 BOE)			0.014 (0.021)		0.005 (0.022)
Constant	6.071*** (0.005)	6.059*** (0.013)	6.052*** (0.013)	6.037*** (0.009)	6.054*** (0.012)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Mon-by-Quint FE	No	Yes	Yes	Yes	Yes
N	48,600	48,600	48,600	49,990	49,990
Firms	4,571	4,571	4,571	4,583	4,583

Notes: This table reports the results of 5 unweighted OLS regressions. The dependent variable is the log of monthly oil production in barrels at the firm level. In Columns (1) – (3), the sample is limited to 2002, and to firms who renewed their annual operating license in that year (*i.e.*, who became bonded) and had positive oil production during 1997–2001. 1[Bonded] is an indicator variable equal to one in all months after a firm’s license renewal month, and equal to zero in all earlier months. Output quintiles are based on average annual production of oil and gas (in barrel-of-oil equivalents) during 1997 – 2001. See text for details. Standard errors are clustered at the operator level. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; \* at the 10% level.

Table 1.6: Effect of Bonding on Natural Gas Production for Remaining Firms

	Implementation Year			Placebo Year	
	(1)	(2)	(3)	(4)	(5)
1[Bonded]	-0.017 (0.012)	-0.007 (0.012)		-0.001 (0.014)	
1[Bonded]*Quintile 1 (1,800 BOE)			-0.027 (0.035)		-0.022 (0.040)
1[Bonded]*Quintile 2 (5,700 BOE)			0.042 (0.032)		-0.007 (0.036)
1[Bonded]*Quintile 3 (16,100 BOE)			-0.028 (0.027)		0.039 (0.036)
1[Bonded]*Quintile 4 (60,200 BOE)			-0.034 (0.028)		-0.017 (0.023)
1[Bonded]*Quintile 5 (60,000,000 BOE)			0.011 (0.017)		-0.003 (0.024)
Constant	7.374*** (0.007)	7.343*** (0.015)	7.341*** (0.015)	7.311*** (0.015)	7.309*** (0.016)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Mon-by-Quint FE	No	Yes	Yes	Yes	Yes
N	41,493	41,493	41,493	43,180	43,180
Firms	4,067	4,067	4,067	4,082	4,082

Notes: This table reports the results of 5 unweighted OLS regressions. The dependent variable is the log of monthly natural gas production in MCF at the firm level. In Columns (1) – (3), the sample is limited to 2002, and to firms who renewed their annual operating license in that year (*i.e.*, who became bonded) and had positive gas production during 1997–2001. 1[Bonded] is an indicator variable equal to one in all months after a firm’s license renewal month, and equal to zero in all earlier months. Output quintiles are based on average annual production of oil and gas (in barrel-of-oil equivalents) during 1997–2001. See text for details. Standard errors are clustered at the operator level. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; \* at the 10% level.

Table 1.7: The Effect of Increased Bond Requirements on Environmental Incidents

	Orphan Wells (OLS)		Water Protection Rules Violations (Negative Binomial)		Well Blowouts (Negative Binomial)	
	(1)	(2)	(3)	(4)	(5)	(6)
1[ <i>After</i> ]	-0.044*** (0.007)	-0.032** (0.015)	-0.148** (0.067)	-0.212 (0.142)	-0.576*** (0.159)	-1.107*** (0.299)
% Change in Expected Count	<b>-56%</b>	<b>-51%</b>	<b>-14%</b>	<b>-19%</b>	<b>-44%</b>	<b>-67%</b>
Constant	0.078*** (0.005)	0.063*** (0.016)	-8.054*** (0.048)	-8.087*** (0.103)	-6.666*** (0.080)	-6.164*** (0.196)
Time Trend	None	Quadratic	None	Exponential	None	Exponential
Month-of- Year FE	No	Yes	No	Yes	No	Yes
Normalization	None	None	Wells Operated	Wells Operated	Wells Drilled	Wells Drilled
N	5,307	5,307	593,891	593,891	56,771	56,771
Firms	5,307	5,307	9,809	9,809	5,043	5,043

Notes: This table reports the results of six separate regressions. In each case, the sample period covers June 1996 to June 2006. Columns (1) and (2) show OLS regressions where the dependent variable is an indicator variable equal to one if the firm orphaned any wells that appear on the March 2014 version of the RRC Orphan Well list. All firms that exited during the sample period are included. There is one observation per firm in its month of exit. Columns (3) and (4) show negative binomial regressions where the dependent variable is the monthly count of violations of Statewide Rules 8 or 14. The sample includes all firm-months with positive oil or gas production. Columns (5) and (6) show negative binomial regressions where the dependent variable is the monthly count of well blowouts. The sample includes all firm-months with drilling activity. In Columns (3)–(6), the dependent variable is normalized by including the log of the normalizing quantity as a regressor with coefficient constrained to one. For OLS, the percentage change in expected count is the regression coefficient ( $\phi$ ) divided by the sample mean prior to the policy change. For negative binomial, the percentage change in expected count is  $e^\phi - 1$ . During the sample period, there were 299 firms that orphaned wells, 5,417 water protection rules violations, and 212 well blowouts. Standard errors are clustered by quarter. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

Table 1.8: The Effect of Increased Bond Requirements: Intensive Margin Effects

	Water Protection Rules Violations (Negative Binomial)		Well Blowouts (Negative Binomial)	
	(1)	(2)	(3)	(4)
1[ <i>After</i> ]	0.056 (0.084)	-0.180** (0.170)	-0.677*** (0.160)	-0.964*** (0.277)
% Change in Expected Count	<b>6%</b>	<b>-16%</b>	<b>-49%</b>	<b>-62%</b>
Constant	-8.388*** (0.072)	-8.223*** (0.155)	-6.602*** (0.089)	-6.297*** (0.244)
Time Trend	None	Exponential	None	Exponential
Month-of- Year FE	No	Yes	No	Yes
Normalization	Wells Operated	Wells Operated	Wells Drilled	Wells Drilled
N	455,566	455,566	44,128	44,128
Firms	5,026	5,026	2,826	2,826

Notes: This table reports results for the specifications in Columns (3) – (6) of Table 1.7, after further limiting the sample to firms that entered the industry prior to December, 2001 and exited after January, 2003. See text for details on each regression. Standard errors are clustered by quarter. \*\*\* indicates statistical significance at the 1% level; \*\* at the 5% level; and \* at the 10% level.

## Chapter 2

# A Credible Approach for Measuring Inframarginal Participation in Energy Efficiency Programs

### 2.1 Introduction<sup>\*</sup>

Global energy consumption is forecast to increase 56% by 2040. While the energy mix is becoming somewhat less carbon-intensive, carbon dioxide emissions are still forecast to increase by 45% over the same period.<sup>1</sup> There is wide agreement among economists that the best policy to reduce carbon dioxide emissions and other negative externalities from energy use would be a Pigouvian tax. Although there has been some recent progress, the vast majority of carbon dioxide emissions worldwide remain untaxed and there are many countries, including the United States, where it seems unlikely that there will be large-scale carbon policy in the near term.

Instead what is receiving much attention is energy efficiency. Electric utilities in the United States, for example, spent \$34 billion on energy-efficiency programs between 1994 and 2012.<sup>2</sup> Energy-efficiency measures like appliance replacement, industrial process changes, and weatherization have the potential to greatly reduce energy consumption (National Academy of Sciences, National Academy of Engineering, and National Research Council, 2010). Proponents of energy-efficiency policies argue that these savings are available at very low cost

---

<sup>\*</sup>The material in this chapter is from co-authored work with Lucas Davis that was published in *Journal of Public Economics*: “A Credible Approach for Measuring Inframarginal Participation in Energy Efficiency Programs”, 2014, vol. 113, issue C, pages 67-79.

<sup>1</sup>These statistics come from U.S. DOE, EIA, International Energy Outlook, released July 2013, Figures 1 and 10. Global energy consumption increased from 350 quadrillion Btu in 1990 to 520 in 2010, and is forecast to increase to 820 by 2040. Energy-related carbon dioxide emissions increased from 20 billion metric tons in 1990 to 30 billion in 2010, and are forecast to increase to 45 billion by 2040.

<sup>2</sup>U.S. Department of Energy, *Electric Power Annual*, 1995-2013. All dollar amounts in the paper are reported in year 2010 dollars. Spending increased every year from 2004 to 2012, with \$4.2 billion in 2012.



(McKinsey and Company, 2009). Thus, energy-efficiency policies are promoted as “win-win” policies that reduce both private energy expenditures and the externalities associated with energy use.

Despite all of the resources aimed at energy-efficiency programs, there is a surprisingly small amount of direct evidence evaluating their effectiveness. A recent review paper emphasizes this lack of evidence and goes on to argue that there is, great potential for a new body of credible empirical work in this area, both because the questions are so important and because there are significant unexploited opportunities for randomized control trials and quasi-experimental designs that have advanced knowledge in other domains. (Allcott and Greenstone, 2012).

We are particularly interested in the question of additionality. Many energy-efficiency programs work by subsidizing households and firms to adopt energy-efficient technologies. A fundamental question in evaluating the cost-effectiveness of these programs is how many of the participants would have adopted these technologies with a lower subsidy, or even with no subsidy at all. Economists have long argued that many participants in energy-efficiency programs may be non-additional or “free riders” (Joskow and Marron, 1992), but demonstrating this empirically has been difficult.<sup>3</sup>

Determining the causal relationship between subsidies and technology adoption is challenging because one must construct a credible counterfactual for adoption in the absence of the policy. Cross-sectional comparisons are misleading because places with generous subsidies are different from places with less generous subsidies. For example, “green” communities like Berkeley, California have more generous subsidy programs but also more eager adopters. Similarly, although programs change over time, it is difficult to separate the causal effect of these changes from other time-varying factors. Changes over time in energy-efficiency subsidies are correlated with changes in technology, pricing, and consumer preferences.

In this paper we address these challenges using a regression discontinuity (RD) analysis. Many energy-efficiency programs have eligibility cutoffs and our paper illustrates how these thresholds can be used to measure inframarginal participation. We apply this approach to a national appliance replacement program in Mexico. We first examine the eligibility thresholds carefully, demonstrating clear discontinuous changes in subsidy amounts and testing for manipulation of the running variable. We then turn to the main analysis, finding that program participation increases noticeably with larger subsidy amounts. For example, when a refrigerator subsidy increases from \$30 to \$110 (both in U.S. 2010 dollars), the number of participants increases by 34%. Thus, the participation elasticity is substantial. However, it is also evident that there are a large number of inframarginal participants. At this threshold, for example, our estimates indicate that about 75% of households would have participated in the program even with the lower subsidy amount. For the four main thresholds in our anal-

---

<sup>3</sup>The term “free rider” has long been used in the context of energy-efficiency programs to describe participants who receive a subsidy for doing something they would have done anyway. This is distinct from the use of the term in economics. The well-known “free rider problem” in economics is that individuals underinvest in public goods because they do not internalize the benefits to others. To avoid confusion we use the term “non-additional” throughout the paper.

ysis we find that 65%+ of households are inframarginal. This large fraction of inframarginal households means that the larger subsidy amounts are almost certainly not cost-effective because each actual increased participant costs a large amount in additional program funds.

We next use the observed changes in demand at these four thresholds to infer what fraction of participants would have participated with no subsidy whatsoever. Under reasonable assumptions, the estimates imply that about half of all participants would have replaced their appliances with no subsidy. We then discuss the implications of non-additionality for cost-effectiveness and welfare. These non-additional participants add cost to the program without yielding any actual reductions in energy consumption. When the marginal cost of public funds is larger than one or when there are indirect program costs then it does not make sense to think of these payments as pure transfers. Our results also demonstrate the potential for cost savings if program designers can target subsidies towards groups where the number of likely non-additional participants is low.

Our paper is the first that we are aware of to use RD to study participation in an energy-efficiency program. We see broad potential for applying this approach in evaluating similar programs. Although eligibility requirements vary widely across programs, the desire to simplify program design often results in the kind of discrete thresholds that we exploit here.<sup>4</sup> In addition, energy consumption is typically carefully measured for large numbers of participants and non-participants. Both of these features make RD a natural approach for causal inference in this context. Relative to the alternative of randomized control trials (RCTs), RD is limited by its focus on specific thresholds. However, RD is easier and less expensive. In addition, RD analyses with administrative datasets have more power and thus can measure smaller effects than typical RCTs.

Most previous studies of additionality in similar programs have been of a much smaller scale (see, e.g., Hartman, 1988), or based on stated-choice experiments (Revelt and Train, 1998; Grosche and Vance, 2009; Bennear, Lee, and Taylor, 2013). Several related papers look at the impact of subsidies on adoption of energy-efficient vehicles (Chandra, Gulati, and Kandlikar, 2010; Gallagher and Muehlegger, 2011; Sallee, 2011; Mian and Sufi, 2012). There is also a small literature which addresses additionality indirectly by comparing realized aggregate savings at the utility level to engineering estimates (Loughran and Kulick, 2004; Auffhammer, Blumstein, and Fowlie, 2008; Arimura et al., 2012). Our paper differs from all of these previous studies because of the RD research design. Probably the closest existing study is Ito, Forthcoming, which uses an RD analysis to examine a California policy that paid households to reduce their electricity consumption in Summer 2005.

---

<sup>4</sup>For instance, the two largest utilities in California offer rebates for energy-efficient heating and cooling equipment that vary across 16 climate zones. These zones were established by California law in 1978 as a function of climate characteristics. Cities can straddle multiple climate zones, and there are large discontinuous changes in rebates at climate zone boundaries. For example, during 2013 Southern California Edison offered three different subsidy amounts (\$550, \$850, and \$1100) for central air conditioners. Other eligibility thresholds that would be amenable to RD analyses include requirements about the vintage of the home, size or characteristics of the households current equipment, and, for needs-based programs, household income.

The paper is also related to a broader literature that examines government programs that subsidize socially-beneficial behavior. A key issue with these programs is the need to distinguish between additional and non-additional participants. Examples include tax subsidies for charitable giving (Feldstein and Clotfelter, 1976), subsidies for building low-income housing (Sinai and Waldfoegel, 2005), conditional cash transfer programs (De Janvry and Sadoulet, 2006), pollution offset programs (Schneider, 2007), and environmental conservation programs (Snchez-Azofeifa et al., 2007).<sup>5</sup>

## 2.2 Conceptual Framework

### Technology Adoption With Externalities

In this section we propose a simple framework for thinking about the costs and benefits of energy-efficiency subsidies. We illustrate the welfare loss introduced by transfers to inframarginal participants and show how the optimal subsidy amount depends on the relative shares of marginal and inframarginal participants. We focus on adoption of an energy-efficient technology, but the same basic framework applies to many other types of government programs that subsidize socially-beneficial behavior.

We begin with a simple graphical partial equilibrium analysis. Figure 2.1 describes the market for an energy-efficient technology. Along the x-axis is the number of adopters. Demand is given by the downward-sloping private marginal benefits curve. The benefits of adoption vary across potential adopters due to differences in expected utilization and other factors. Supply is described by the private marginal cost curve.

The privately optimal level of adoption is labeled in the figure as  $Q_0$ . These consumers adopt the technology purely on the basis of private benefits, even with no subsidy or other form of government intervention. If there are no externalities, then  $Q_0$  is socially optimal. Once externalities are introduced, however, this is no longer the case. The figure illustrates the case in which there is a positive marginal external benefit from adoption, so the social marginal benefit exceeds private marginal benefit. The socially optimal level of adoption is labeled in the figure as  $Q^*$ . This optimum is defined as the intersection of the social marginal benefit and private marginal cost curves. The optimal subsidy is  $s^*$ . With this subsidy, adopters between  $Q_0$  and  $Q^*$  are additional. They adopt under the subsidy and do not adopt without it.

The total amount paid in subsidies is indicated by the rectangle A/B/C. Rectangle A is a transfer to non-additional participants, i.e. consumers who would have adopted the energy-efficient technology even with no subsidy whatsoever. Triangle B is excess payment

---

<sup>5</sup>In this broader literature there are a few studies that use RD. Baum-Snow and Marion (2009) examines the effect of tax credits for building low-income housing, exploiting a discontinuous increase in the credit amount in census tracts where more than 50% of households qualify for means-tested government housing assistance. Filmer and Schady (2011) studies a conditional cash transfer program in Cambodia where program eligibility is limited to households scoring below a specified level on a government poverty index.

to consumers who are induced to adopt because of the subsidy. Most adopters receive a subsidy that is more than the minimum amount necessary to induce them to adopt. And triangle C is the payment required to make adopters between  $Q_0$  and  $Q^*$  indifferent between adopting and not adopting.

Before proceeding it is worth highlighting a couple of important assumptions. First, we have assumed that the external benefits from adoption are the same for all potential adopters. When external benefits differ there can be gains from targeting energy conservation policies toward high value participants (Allcott, 2015; Allcott, Mullainathan, and Taubinsky, 2014). We have also assumed constant marginal costs. With increasing marginal costs the analysis is similar but the incidence of the subsidy is partly on sellers. As a result there are “non-additional recipients” on both sides of the market. Subsidies increase the equilibrium price of the good, leading to higher revenues for sellers even for transactions which would have occurred anyway.

## Incorporating Pre-Existing Taxes and Other Distortions

This partial equilibrium analysis ignores interactions with taxes and other pre-existing distortions. Consider the following welfare function,

$$W = U(Q(s)) - C(Q(s)) + \tau Q(s) + Q(s)s - \eta Q(s)s. \quad (2.1)$$

Here  $Q(s)$  is the quantity of technology adoption, which is a weakly increasing function of the subsidy  $s$ .  $U(\cdot)$  and  $C(\cdot)$  are private benefits and costs from the energy-efficient technology. In the graphical analysis, these correspond to the areas under the private marginal benefit and private marginal cost curves to the left of  $Q$ .  $\tau$  is the constant external benefit of technology adoption derived from, for example, reduced carbon dioxide emissions.

The final two terms reflect general equilibrium effects. The subsidy payments are a transfer from taxpayers to adopters in the amount  $Q(s)s$ . The efficiency cost of interactions with pre-existing distortions is denoted  $\eta$ . If  $\eta$  is one, then there is no efficiency loss associated with the transfers, and the gains by adopters exactly offset the costs to taxpayers.

The welfare change from a marginal increase in the subsidy is given by,

$$\frac{dQ}{ds} \left[ U'(Q(s)) - C'(Q(s)) + \tau - (\eta - 1)s \right] - (\eta - 1)Q(s). \quad (2.2)$$

The additional adoption induced by the subsidy increase is  $\frac{dQ}{ds}$ . The left-hand term gives the welfare effect of bringing these marginal participants into the program: private marginal benefits minus private marginal costs, plus external benefits, minus the efficiency cost of financing the subsidy payments to new participants. The right-hand term gives the welfare cost of increased payments to inframarginal participants: The  $Q(s)$  participants already adopting the technology each receive an infinitesimal increase in subsidy payment, financed at a cost of  $\eta$ . The welfare effects of increasing the subsidy depend on the relative numbers of marginal and inframarginal participants. If  $\frac{dQ}{ds}$  is large relative to  $Q(s)$ , then the left-hand

term matters more than the right-hand term. As the subsidy level increases,  $Q(s)$  becomes larger and payments to inframarginal participants become more and more important.

If  $\eta$  is equal to one then equation (2.2) simplifies considerably and it is optimal to set the subsidy equal to marginal external benefits ( $\tau$ ). This is exactly what we described in Figure 2.1 with  $s^*$  and  $Q^*$ . However, if  $\eta$  is greater than one then the optimal subsidy level is below marginal external benefits. The optimal subsidy amount balances the benefits of increased adoption with the full welfare costs, including the general equilibrium efficiency costs of larger transfers.

The value of  $\eta$  is informed by a large literature on the general equilibrium effects of environmental taxes and subsidies, which we quickly summarize here. See, e.g., Bovenberg and Goulder, 2002 and references therein. These policies create “tax interaction” and “revenue recycling” effects. Environmental taxes exacerbate pre-existing distortions in the economy, for example, by further decreasing the real wage in the presence of a labor tax (the tax interaction effect). At the same time, environmental taxes also generate revenues, allowing labor and other distortionary taxes to be lower than they would be otherwise (the revenue recycling effect). A series of analytical and numerical studies have concluded that, for taxes, tax interaction is more important than revenue recycling, so that optimal tax rates on externalities are generally below marginal damages (Bovenberg and Goulder, 1996; Bovenberg and Goulder, 2002). One reason for this is that environmental taxes discourage consumption of the taxed good, which erodes the tax base and undermines revenue recycling.

A symmetric set of results holds for environmental subsidies. Subsidies for “clean” goods increase real wages, thereby *decreasing* the distortionary effects of labor taxes (the tax interaction effect). However, subsidies also require labor and other taxes to be *higher* than they would be otherwise, exacerbating distortions (the “revenue-financing” effect). In our framework,  $\eta$  represents the net of these two effects. The literature suggests that, for subsidies, the revenue-financing effect exceeds the tax-interaction effect (Parry, 1998). Thus,  $\eta$  is greater than one and the optimal level of subsidy is positive, but below marginal external benefit, just like the optimal level of tax is positive, but below marginal external damages. Throughout the paper, we refer to  $\eta$  as the efficiency cost of transfers.

The preceding analysis assumes that the government must pay all adopters the same subsidy. Targeting subsidies to different groups based on expected benefits and costs could increase welfare substantially. In the extreme, perfect price discrimination would pay each individual adopter the minimum amount they require to adopt. There would be no payments to inframarginal participants and the right-hand term in equation (2.2) would disappear. However, in practice, equity concerns, imperfect information, and other factors prevent perfect price discrimination and limit group-level targeting.

The main takeaway from this section is that, in general, the optimal subsidy amount is lower than marginal external benefits. The welfare effects of a subsidy depend critically on the effect of the subsidy on program participation. If  $\frac{dQ}{ds}$  is small relative to the number of existing participants, then the benefits from increased adoption will be small relative to the efficiency costs of the payments to inframarginal participants. Accordingly, this is where we focus our attention in the empirical analyses which follow.

## 2.3 Program Background and Construction of Dataset

### Background

Our empirical analysis focuses on a large-scale energy-efficiency program in Mexico. The program was launched in March 2009 and ended in December 2012. During this period, the program subsidized the replacement of 1.9 million refrigerators and air conditioners with energy-efficient models. Davis, Fuchs, and Gertler (2014) compare electricity consumption by program participants before and after appliance replacement, finding that realized savings were considerably smaller than what was predicted by *ex ante* analyses. There was no attempt in this previous work, however, to distinguish between additional and non-additional participants, nor was there any examination of the programs eligibility thresholds or RD analysis.

To participate in the program a household had to have a refrigerator or air conditioner that was at least 10 years old and agree to purchase a new appliance meeting Mexican energy-efficiency standards. The old appliances were transported to recycling facilities and disassembled. The refrigerator subsidies were available nationwide. For the air conditioner program, a household needed to live in one of four officially-designated climate zones with a mean summer temperature of at least 30°C (86°F); this included about one-quarter of all households.

Table 2.1 describes the subsidies available under the program. The direct cash payments came in three different amounts, approximately corresponding to \$30, \$110, and \$170 (all in U.S. 2010 dollars). Eligibility for the different subsidy amounts depended on a household's average historical electricity consumption, calculated over the previous year. There was a minimum consumption level below which households were ineligible for subsidies. Above this minimum, the cash payment amount decreased with a household's consumption level. This structure was designed to target the larger subsidies to lower-income households.

The program also offered on-bill financing at an annual interest rate of 13.8%, repaid over four years. At the first two thresholds, the increase in maximum loan amount exactly equals the decrease in cash. If households would otherwise have financed the purchases using credit cards, the increase in maximum loan amount offsets about 18% of the decrease in cash subsidies at these thresholds.<sup>6</sup> At the highest consumption threshold, the increase in maximum loan amount greatly exceeds the decrease in direct cash payments. For households with a typical cost of borrowing who take the full loan amount, the economic value of the program actually increases at this threshold.

---

<sup>6</sup>According to Banco de Mexico, Indicadores Básicos de Tarjeta de Crédito, October 2012, Table 1, credit cards in Mexico charged an average interest rate of 25.3% in 2011. This is not a perfect measure. On the one hand, not all households have access to credit cards, and the interest rates on other forms of borrowing will vary. On the other hand, collateralized loans for durable goods purchases typically can be made at lower rates. Most participants took out at least some loans, suggesting that the market cost of borrowing exceeds 13.8% in most cases.

There are several features of this program that make it particularly conducive to an empirical analysis. First, retailers did not have any discretion in assigning subsidy amounts. Participating retailers determined which subsidy a household was eligible for by entering the household's account number into a website designed for this purpose, and this online record became part of the paperwork necessary for the retailer to be reimbursed. This lack of scope for retailer discretion is important because even a small amount of selection at these thresholds would have been a threat to our identification strategy.

Another nice feature of the program is that participants received these subsidies immediately. In order to participate, a household was required to show a recent electricity bill and an identification card, but there was no paperwork required and no delay in receiving the subsidy. This differs from appliance subsidy programs in the United States which typically require participants to fill out and mail application forms and proofs of purchase, and then wait for a rebate check to arrive in the mail. In programs for which there are "hassle" costs like these, not all eligible households will participate. And the amount of selection depends on the size of the subsidy, making it difficult to interpret differences in participation across subsidy levels.

## Construction of the Dataset

A key feature of our analysis is the use of high-quality, household-level microdata, both about program participants and about the entire pool of potential participants. The fact that we observe eligibility for non-participants is important because, as usual, the objective in the empirical analysis is to construct a credible counterfactual, and this is hard to do without information about the broader pool.

The first component of this database is a two-year panel dataset of household-level electric billing records describing bimonthly electricity consumption for the universe of Mexican residential customers from May 2009 through April 2011. The complete set of billing records includes data from 25,786,609 households. This represents the entire pool of potential participants in the program.

The second component of this database is a record of all households who participated in the program between March 2009 and June 2011. In the complete dataset there are a total of 1,162,775 participants. We merged this list with the electric billing records using customer account numbers. We used our database to calculate average historical electricity consumption for each household according to the program rules. For details see the Appendix.

We focus on the 237,552 participants in 2011 because calculating average historical electricity consumption for earlier participants would require data from before May 2009, the first month in our billing records. For each participant, we know the exact dates of purchase and replacement, whether the appliance was a refrigerator or an air conditioner, and the amount of direct cash payment and loan received.

## 2.4 Empirical Strategy

### Estimating Equation

Our empirical strategy exploits the discrete eligibility thresholds that determined whether a household was eligible for zero subsidy, \$30, \$110, or \$170. There are six total thresholds; three for air conditioners and three for refrigerators. At each of these thresholds, we use a standard RD estimating equation (Lee and Lemieux, 2010):

$$1[Participate]_i = \alpha + f(X_i) + \rho 1[Below Threshold]_i + \eta_i \quad (2.3)$$

where  $1[Participate]_i$  is an indicator variable equal to one if a household participated in the program and zero otherwise. We include in the regression  $f(X_i)$ , a polynomial in average historical electricity consumption, and  $1[Below Threshold]_i$  an indicator variable equal to one if the household's average historical electricity consumption was below the given threshold. The coefficient of interest is  $\rho$ , which measures the discontinuous change in program participation at the threshold. Moreover, we normalize  $X_i$  to be equal to zero at the threshold so the coefficient  $\alpha$  corresponds to the predicted probability of participating just below the threshold, and  $\alpha + \rho$  corresponds to the predicted probability just above the threshold. In terms of the conceptual framework described in Section 2.2,  $\rho$  is the empirical analog of  $\frac{dQ}{ds}$ , and  $\alpha$  is the empirical analog of  $Q(s)$ .

The error term  $\eta_i$  captures unobserved determinants of the participation decision. An important advantage of RD is that it requires a considerably weaker identifying assumption than other approaches. Hahn, Todd, and Klaauw, 2001 show that identification with RD requires that the conditional mean function  $E[\eta_i|X_i]$  is continuous at the discontinuity. In the limit, one is comparing outcomes within an arbitrarily small neighborhood around each threshold and the identifying assumption requires only that there not be a discontinuous change in these other factors that occurs exactly at the eligibility thresholds. Of course, in practice there are few observations within an arbitrarily small neighborhood around these thresholds, and so there is a trade-off between bias and efficiency. Flexibly parameterizing the polynomial  $f(X_i)$ , allows us to expand the sample to include households farther away from the threshold.

We report results using several different bandwidths. In our preferred specification, we include all households within 100 kilowatt-hours of the thresholds for air conditioners, and within 50 kilowatt-hours of the thresholds for refrigerators. The wider bandwidth for air conditioners reflects that these thresholds were much higher (500, 750, and 1000 kilowatt-hours compared to 175, 200, and 250) and the density of households in that part of the distribution is lower. With refrigerators, the thresholds are close enough together that, in some cases, the bandwidth includes more than one threshold. In the results which follow we use one estimating equation per threshold, but we include intercept terms for any additional thresholds.



## Validity of Research Design

### The Discontinuity in Subsidy Amounts

Figure 3.4 plots the fraction of participants who received the larger subsidy as a function of average historical electricity consumption. The dots represent mean values for three kilowatt-hour bins. In all six cases there is a clear discontinuity at the threshold. Almost all households with average historical consumption below the threshold receive the higher subsidy amount and almost all households with average historical consumption above the threshold receive the lower subsidy amount. Figure 3.4A is typical of all three air conditioner thresholds. The share of participants receiving the larger subsidy falls from near one to near zero. Even within very narrow bandwidths around these thresholds, we are able to correctly predict subsidy levels for 99%+ of all participants (see the Appendix for details).

The discontinuities are less sharp for refrigerators. Figure 3.4D is typical of the three refrigerator thresholds. Near the threshold, a small number of participants receive a different subsidy than we would have predicted. This is due to measurement error in our reconstruction of average historical electricity consumption. As we explain in more detail in the Appendix, the program rules for refrigerators were especially complicated, introducing a small amount of measurement error for some observations. Battistin et al. (2009) show that this type of measurement error biases sharp RD estimates downward in proportion to the fraction of observations measured with error, but that the fuzzy RD estimator is unbiased as long as the measurement error is uncorrelated with the subsidy amount. In practice, because a small share of observations are measured with error, sharp RD and fuzzy RD produce very similar estimates.

### Checking for Manipulation of the Running Variable

A standard concern with RD analyses is manipulation of the running variable. If participants could completely or partially manipulate their treatment status, this would represent a substantial threat to the identifying assumption. Understanding any strategic behavior in response to eligibility thresholds is also of significant independent interest because it may introduce inefficiencies, as agents alter their behavior to qualify for more generous subsidies (Sallee and Slemrod, 2012).

Figure 2.3 plots the frequency distribution of average historical electricity consumption for all households. We use three kilowatt-hour bins and include separate plots for air conditioners and refrigerators because the measure of average historical electricity consumption used to determine eligibility was different for the two appliance types. Examining the smoothness of the running variable is a valuable first test for manipulation (McCrary, 2008). If households were changing their behavior to qualify for the more generous subsidy, we would expect to see bunching to the left of the thresholds. For both appliance types, the frequency distributions appear smooth across all eligibility thresholds. This lack of evidence of manipulation is perhaps not surprising given that it is difficult for a household to control its average historical electricity consumption. Perhaps most importantly, this is *historical* consumption, so at the

time of participating in the program, there is no scope for the household to go back and change its electricity consumption patterns in the past.

Another standard RD specification test is to look for changes at the threshold in covariates unrelated to the treatment variable. If manipulation of the running variable leads to systematic sorting of households around the threshold, we would expect to see discontinuous differences in household characteristics at the threshold. In our dataset, we do not have any household-level covariates. Instead, we merged our dataset with municipality-average household income from the 2010 Census. Figure 2.4 shows that municipality-average income is smooth across all eligibility thresholds, suggesting that there is no discontinuous change at the threshold in the affluence of the places where participants live.

Finally, we consider a more subtle form of strategic behavior. If a household somehow learned that they just missed qualifying for a larger subsidy, they could in theory wait one or more billing cycles, perhaps while intentionally reducing electricity consumption, and then reapply. In practice the program structure made this unlikely.<sup>7</sup> Moreover, in the Appendix we test for this explicitly by comparing participants' historic average electricity consumption in the months before participation to nonparticipants' historic average electricity consumption over the same months. We find no evidence that participants were more likely than nonparticipants to become eligible for larger subsidies immediately before participating. Thus, there is no evidence of strategic delay.

## 2.5 Results

### Graphical Evidence

We now turn to our main results, first presenting graphical evidence and then reporting regression estimates in Section 2.5 and alternative specifications in Section 2.5. Figures 2.5A and 2.5B plot program participation against average historical electricity consumption for air conditioners and refrigerators, respectively. We again use three kilowatt-hour usage bins. The y-axis in these figures is the percentage of households in each bin that participated in the program during our sample period. For refrigerators the denominator is all Mexican households. For air-conditioners the denominator is all Mexican households living in climate zones that were eligible for the air conditioner program.

It is first worth noting that there is essentially no participation by households who used less than the minimum levels of electricity required for participation. This is reassuring, though not surprising given the way the program was administered. The small number of participating households to the left of the minimum eligibility thresholds for refrigerators reflects a small amount of measurement error in average historical electricity consumption.

---

<sup>7</sup>Participating retailers determined whether a household was eligible by entering the household's account number into a website designed for this purpose. Households could not access this site without a retailer's login and password. The website reported the subsidy level for which a household qualified, but did not describe the intermediate calculations which determined eligibility or let a household know when it was close to a more generous subsidy level.

For air conditioners, participation increases steadily between 250 and 500 kilowatt-hours, levels off between 500 and 750, and then declines slowly after 750. Our main interest is in behavior at the 500, 750, and 1000 kilowatt-hour thresholds. In the first two cases, there appears to be a discontinuous decrease in participation at the threshold. The second decrease is particularly visible and appears to occur exactly at the threshold in which the subsidy amount decreases from \$110 to \$30. It is difficult to make strong statements based on this graphical evidence because the participation rate moves around across bins, but at this threshold the participation rate appears to drop from about 1.5% to about 1%. At the final threshold, where the cash subsidy amount falls from \$30 to zero, there does not appear to be any discontinuous change in participation.

For refrigerators, participation follows a similar inverted “U” pattern, peaking at about 1.8% near 150 kilowatt-hours and then decreasing steadily between 150 and 300. At both the 175 and 200 kilowatt-hour thresholds there are visible discontinuous decreases in participation. At the 250 kilowatt-hour threshold there is no apparent decrease. This general pattern is similar to what is observed for air conditioners, with decreases at the first two thresholds and no visible decrease at the third threshold.

For both appliance types, there is no observed change in participation when the subsidy falls from \$30 to \$0. As we discussed in Section 2.3, this threshold was different from the others in that there was a large offsetting increase in the maximum loan amount. We were expecting to see a much smaller change in participation at this threshold, and the data appear to bear this out. The near zero change in participation implies that, on average, the increase in maximum loan amount had about the same value to households as the \$30 decrease in cash. We find this very interesting, but in the regression analysis which follows we focus on the four other thresholds where there is a clear and unambiguous change in the value of the program.

## Regression Estimates

Table 2.2 reports RD estimates and standard errors from four separate regressions. For each threshold we report the percentage of households participating at each side of the threshold as well as the percent change in participation. Because we have normalized the running variable to be equal to zero at the threshold, these statistics come right out of our estimating equation. From each regression, column (2) reports the estimated intercept, column (3) reports our estimate of the intercept plus our estimate of the discontinuous change at the threshold, and column (4) reports the percent change between the two. Columns (5) and (6) report the implied linear slope of demand and price elasticity at each threshold.

Consistent with the graphical evidence, participation increases at all four thresholds. All four changes are statistically significant (three at the 1% level, one at the 5% level). With air conditioners, the increases are 21% and 45%. For refrigerators the estimated changes in participation are similar, 19% and 34%. As with air conditioners, the larger increase corresponds to the subsidy increase from \$30 to \$110. The estimates for refrigerators are

more precisely estimated because of the large number of households with average historical electricity consumption near these thresholds.

The increases are clear, but the estimates also imply that a large number of participants are inframarginal. Most households who just barely qualified for the \$170 subsidy would have participated even if they had only received \$110 and most households who just barely qualified for the \$110 subsidy would have participated even if they had only received \$30. The percent inframarginal can be calculated by dividing column (2) by column (3). For example, when the air conditioner subsidy increases from \$110 to \$170, our estimates imply that 83% ( $\frac{1.45}{1.75} = 0.83$ ) of households are inframarginal. Across thresholds the percentage inframarginal ranges from 69% to 84%. The estimates are similar for air conditioners and refrigerators, suggesting that the results are not driven by idiosyncratic features of a particular appliance market.

In column (5) we report the implied slope of demand at the threshold. These are calculated for each threshold by dividing the percent change in participation by the subsidy change in dollars. For each \$1 of subsidy change, the share of households replacing air conditioners increases 0.0054 to 0.0069 percentage points and the share of households replacing refrigerators increases 0.0044 to 0.0047.<sup>8</sup> These slopes appear quite small but it is important to keep in mind that the base participation rates are very low.

In column (6) we report the implied price elasticities. These are calculated for each threshold by dividing the percent change in participation by the percent change in the price of appliance replacement net of the subsidy. In calculating this net price we use the average appliance price paid by program participants at the threshold.<sup>9</sup> For air conditioners, the elasticity is 0.88 at the first threshold and 1.76 at the second. For refrigerators, the elasticity is 0.90 at the first threshold and 1.51 at the second.<sup>10</sup>

It is important to interpret these elasticities carefully. They describe how demand would change in response to a market-wide price change, not the elasticity of demand for a particular appliance model or for all appliances made by a particular manufacturer. It is also worth emphasizing that this is the elasticity of demand for appliance *replacement*, which is different from demand for initial purchase. Still, the estimates appear quite large. They imply that

---

<sup>8</sup>For the calculations in columns (5) and (6) we calculate the change in the value of the subsidy incorporating both direct cash payments and the implied cash value of the on-bill financing (assuming a 25.3% annual interest rate on private borrowing; see Section 2.3).

<sup>9</sup>Specifically, we use the average price paid by participants at the low-subsidy side of each threshold. For air conditioners, these prices were \$402 at the 175 kWh threshold and \$402 at the 200 kWh threshold. For refrigerators the prices were \$425 at the 500 kWh threshold and \$427 at the 750 kWh threshold. We calculate all elasticities as arc elasticities, and thus use for the denominator in these calculations the midpoint between the high-subsidy and low-subsidy prices.

<sup>10</sup>We are reporting uncompensated elasticities, but compensated elasticities are likely to be very similar. These subsidies represent a tiny share of the total household budget for these households so income effects are likely negligible. One approach for assessing the potential magnitude of income effects is to test for changes at the thresholds in the price of the appliance that is purchased. We observe no significant change in the price of the appliance purchased at three of the four thresholds. At the fourth, the average price increases by about 2%, and the increase is only weakly statistically significant.

appliance replacement is price-responsive, and that the program caused a large number of appliance replacements that otherwise would not have happened.

These estimates are valuable not only in assessing energy-efficiency subsidies, but also for predicting appliance replacement more broadly. Wolfram., Shelef, and Gertler (2012) argues that demand for residential appliances will have an enormous influence on future energy consumption growth in low- and middle-income countries. Appliance prices have been falling for decades and our estimates imply that continued decreases will accelerate the rate at which appliances are replaced. If households are more quickly replacing appliances this means that improvements in energy-efficiency will more quickly be reflected in the appliance stock.

## Alternative Specifications

Table 2.3 reports regression estimates from five alternative specifications. For each specification we report the estimated percent change in participation at each threshold. In the first three columns, we vary the size of the bandwidth used with the cubic polynomial. The second column reports our baseline estimates, identical to the estimates reported in Table 2.2. The first and third columns assess the sensitivity of our estimates to larger and smaller bandwidths. In the fourth column, we use local linear regression with a uniform kernel and a small bandwidth. Overall, the results are similar across all four columns. Moreover, there is no consistent pattern. As we move across bandwidths and specifications, some point estimates increase while others decrease.

The last column reports estimates from a fuzzy RD specification. In this specification, we scale the estimates by the size of the discontinuity at the threshold following Hahn, Todd, and Klaauw, 2001 and Battistin et al., 2009. Specifically, we run a first stage regression of an indicator for the larger subsidy ( $1[Larger\ Subsidy]$ ) on ( $1[BelowThreshold]$ ) and a cubic polynomial of average historical consumption,  $g(X)$ ,

$$1[Larger\ Subsidy]_i = \phi + g(X_i) + \gamma 1[Below\ Threshold]_i + \epsilon_i. \quad (2.4)$$

We then divide our baseline estimates by  $\gamma$  to remove any bias caused by measurement error (see section 2.4 and the Appendix). The estimates are very similar with the fuzzy RD specification. The air conditioner estimates are essentially identical to the sharp RD estimates, consistent with the near perfect discontinuity observed in Figures 3.4A, 3.4B, and 3.4C. For refrigerators, the scaling increases the point estimates modestly, consistent with the graphical evidence in Figures 3.4D, 3.4E, and 3.4F, which exhibit a somewhat less perfect discontinuity.

## 2.6 Discussion

### Inferring the Fraction Non-Additional

These estimates are directly relevant for program design because they show how adjustments in program generosity would have changed participation levels. We are also interested in what program participation would have been with no subsidy whatsoever. Table 5 reports estimates of the fraction of participants that are non-additional under two different assumptions about the shape of the demand curve.

In Column (1) we calculate the fraction of participating households who are non-additional by using the slope estimates from Table 2.2 to predict appliance replacement at the unsubsidized price. For participants who received the \$170 cash payment, we use the slope corresponding to the threshold between \$110 and \$170, and for participants who received \$30 or \$110, we use the slope corresponding to the threshold between \$30 and \$110. The implied slopes are quite similar across thresholds, however, so the results are not particularly sensitive to which slope we use. Under these assumptions our estimates imply that 54% of participants were non-additional, in that they would have replaced their appliances even with no subsidy whatsoever. With this level of non-additionality, the average payment amount per induced replacement is \$328, a little more than twice the average subsidy amount (\$152).

In Column (2), we infer the fraction of participants that are non-additional by using the estimated *elasticities*, rather than the estimated slopes. With this approach our estimates imply that 43% of participants are non-additional, so that the average payment per induced replacement is \$269. Of the two alternatives we prefer to use the estimated slopes because this assumption about the demand curve better fits the observed behavior at the thresholds. Whereas the estimated slopes are similar across thresholds, the estimated elasticities are not. For both appliance types the estimated elasticities are considerably larger at the \$30 to \$110 threshold than at the \$110 to \$170 threshold, which is what one would expect with linear demand.

It is worth emphasizing that both of these approaches rely on strong assumptions about demand. By using our estimates of these slopes and elasticities to predict behavior away from the thresholds, we are assuming that behavior at the thresholds is representative of all households. This is a mild assumption for participants who are close to thresholds but is a considerably stronger assumption for participants far away from a threshold like those at the beginning of the \$170 tier. We find it somewhat reassuring that the slope estimates are similar across thresholds. Nonetheless, these projections should be viewed with more caution than the RD estimates from which they are derived.

### Implications for Cost-Effectiveness and Welfare

Depending on whether one uses the slopes or the elasticities, we find that 43 – 54% of participants are non-additional. So it costs on average \$269 to \$328 in subsidies per induced replacement, instead of \$152 in a naive analysis that treats all participants as additional.

Thus, accounting for non-additional participants approximately doubles the program cost per unit of reduced energy consumption. Related measures of cost-effectiveness such as the program cost per ton of carbon dioxide abated would also approximately double.

Non-additionality also affects the full welfare calculation. In this section we provide a brief sketch of such a calculation, following the framework outlined in Section 2.2. To calculate the benefits of appliance replacement, we value the reductions in greenhouse gases and local air pollutants from each appliance replacement. As we explain in detail in the Appendix, we use the pre-program engineering estimates of 2,900 kilowatt-hours in lifetime electricity savings per replacement. This is equivalent to 1.6 tons of avoided carbon dioxide emissions so applying a \$34 social cost of carbon as in U.S. IAWG (2013), the climate benefits are \$53 per replacement. We also include \$54 per replacement of benefits from reduced local air pollution.

The program costs can be divided into categories A, B, and C, as indicated in Figure 2.1. Rectangle A represents payments to non-additional participants. Private costs are zero for these participants, since they are doing something they would have done anyway. But there is still the efficiency cost associated with financing the subsidies, which we called  $\eta$  in Section 2.2. This parameter  $\eta$  represents the net welfare cost of the revenue financing and tax interaction effects. For this simple back of the envelope calculation, we assume that  $\eta$  equals 1.3, following Goulder, Parry, and Burtraw (1997) and other studies in the literature (see the Appendix for details). We find that about half of the participants are non-additional and the average subsidy amount was \$152. So, transferring funds to non-additional participants imposed an efficiency cost of about \$46 per induced replacement.

Rectangle BC represents payments to induced participants. Financing these payments cost an additional \$46 per replacement, again using 1.3 for the efficiency cost. In addition, there are the private costs of adoption. These costs are shown as Triangle C. If demand is linear, then this area is half the total amount of subsidies paid to induced participants. The average subsidy amount was \$152, so the private cost of replacement averaged \$76. Summing up these back-of-the-envelope benefits and costs, we find that each induced replacement yielded benefits of \$107 at a cost of \$168.

These calculations highlight the importance of distinguishing between additional and non-additional participants. This can be seen most starkly by comparing these numbers to what one would have calculated with a naive analysis that assumes all participants are additional. In the naive analysis, the efficiency cost of financing the program is much smaller: \$46 per replacement rather than \$92. Private costs are the same (\$76), so the total cost is \$122 per replacement. This is much closer to the benefits of \$107. Thus, in the naive analysis, the program appears much closer to welfare-improving.

These values should be interpreted carefully because they are based on many strong assumptions. These calculations also ignore some components of benefits and costs, such as the benefits of properly disposing of refrigerants and the administrative costs of the program. They also rely on engineering estimates of electricity savings, which recent work suggest may have been overly generous (Davis, Fuchs, and Gertler, 2014). The goal of this simple back-of-the-envelope calculation is to tie our empirical results to the economic model and provide

an example for how to think about welfare analysis in this setting.

Two of the most important uncertain parameters are the social cost of carbon (SCC) and the efficiency cost of financing the subsidies ( $\eta$ ). The program becomes more attractive with a high SCC and low  $\eta$ . Still, it would have taken values near the extremes of the range of available estimates in the literature in order to make the program welfare-improving. Holding constant our other assumptions, the program benefits would exceed the costs only if the SCC were greater than \$73 per ton or if  $\eta$  were less than 1.1.<sup>11</sup>

These results raise questions about whether energy-efficiency programs could be designed differently to target payments based on expected additionality. For example, if immutable, verifiable household and firm characteristics could be determined to predict adoption with and without the subsidy, payments could be made conditional on these characteristics.<sup>12</sup> The scope for this type of targeting will differ widely across contexts and there are important constraints that may limit targeting in practice. Income-based targeting, for example, can be difficult and expensive to enforce, and geographic targeting may be unacceptable politically. Still, even in programs where explicit targeting is limited, softer versions may be possible. As a simple example, perhaps program advertisements can be tailored towards demographic segments where adoption in the absence of the program would be low.

## Program-Wide Effects

RD is well-suited for highly-localized predictions about how participation would have changed under alternative subsidies. But we have also stressed that RD is not a panacea and cannot answer all of the questions that could be answered, for example, with a large-scale RCT. A particularly important weakness is the inability of RD to measure broader program-wide effects. In providing subsidies the government is providing information and an explicit endorsement of particular energy-efficient technologies. This focuses attention on these products, potentially influencing replacement decisions above and beyond the direct impact of the subsidies themselves.

Through program-wide effects even non-participants may have their behavior influenced by a program. For example, potential participants may investigate a program only to learn that they are ineligible for a subsidy. However, in learning about the program they focus their attention on energy-efficiency, potentially becoming more likely to adopt the subsidized technology even if they do not end up receiving any monetary incentive whatsoever.

---

<sup>11</sup>An SCC of \$73 per ton is above the central range of values presented in U.S. IAWG (2013), but less than the 95th percentile estimate of \$129. For  $\eta$  to be less than 1.1, the MCPF would have to be at the bottom of the range of estimates of 1.11 to 1.56 for the United States (Bovenberg and Goulder, 2002), or the tax interaction effect would have to be large.

<sup>12</sup>De Janvry and Sadoulet (2006) propose such targeting for a conditional cash transfer program in Mexico. Using experimental data, they conclude that making school attendance subsidies a function of child's gender, birth order, and distance traveled to school could decrease the program cost per additional child attending school by about 23%.



These broader program impacts are difficult to measure empirically. Ideally, one would measure program-wide effects using a large-scale RCT in which randomization was done not over households, but over geographic areas with subsidy generosity varied across areas. This has been done with cash transfer programs (Baird et al., 2012), and given sufficient resources and public cooperation could be implemented with energy-efficiency programs. Experiments could also be designed to directly measure spillovers through social networks, as in Miguel and Kremer (2004) and similar studies.

## 2.7 Conclusion

It is hard to provide incentives for socially-beneficial behavior without substantial transfers to those who would have done these behaviors anyway. Subsidies for energy efficiency are a key example, both because the potential external benefits are large and because first-best policies seem, for the moment, to be impossible politically. Empirical estimates of additionality are critical, however, because if a large enough fraction of participants are non-additional then a program will not be welfare improving.

Our RD analysis avoids many of the measurement and endogeneity problems in previous studies by focusing on behavior within narrow windows around eligibility thresholds. Although these thresholds make RD a natural approach to causal inference, we are not aware of any previous RD analyses of additionality in this context. We see broad potential for applying our conceptual framework, estimating equations, and tests of strategic behavior in evaluating similar programs. Although the exact eligibility requirements vary across programs, it is typical to see discontinuous thresholds of the type observed here.

The results are striking. We find that most households would have participated even for much lower subsidy amounts. Across thresholds, more than two-thirds of participants are inframarginal and the estimates imply that about half of all participants would have replaced their appliances even with no subsidy whatsoever. These non-additional participants add substantial cost to the program without yielding any real reduction in energy use.

These findings are relevant to current energy policy around the world, which is focusing increasingly on energy efficiency. Billions of dollars are spent each year on programs like this one that provide subsidies for households and firms who adopt energy-efficient technologies. Reliable empirical estimates of the benefits and costs of these policies are essential and using RD to measure changes in behavior at eligibility thresholds can be an important part of these analyses.

Figure 2.1: The Market for an Energy-Efficient Technology

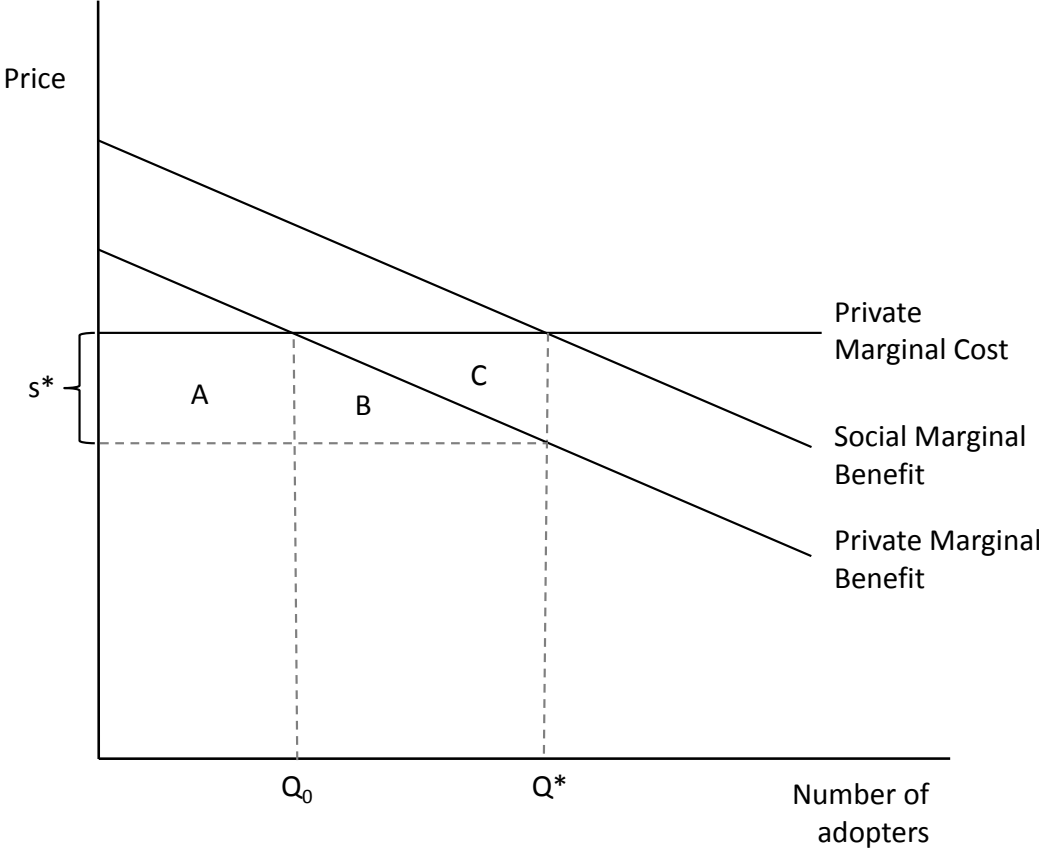
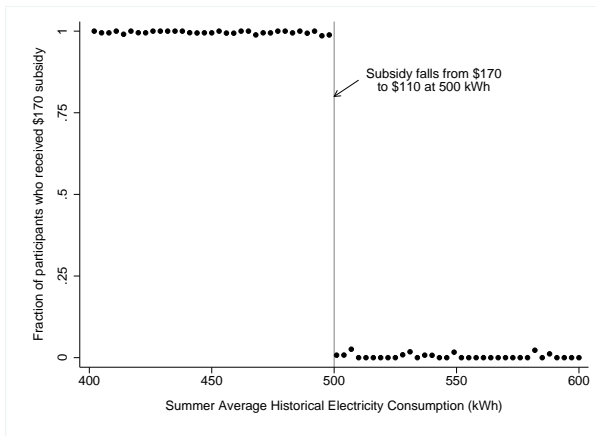
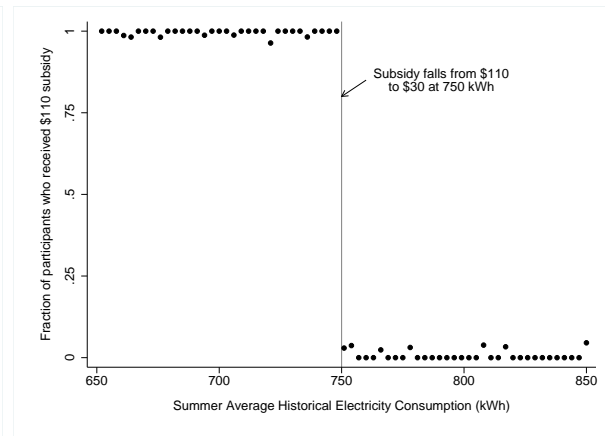


Figure 2.2: The Discontinuities

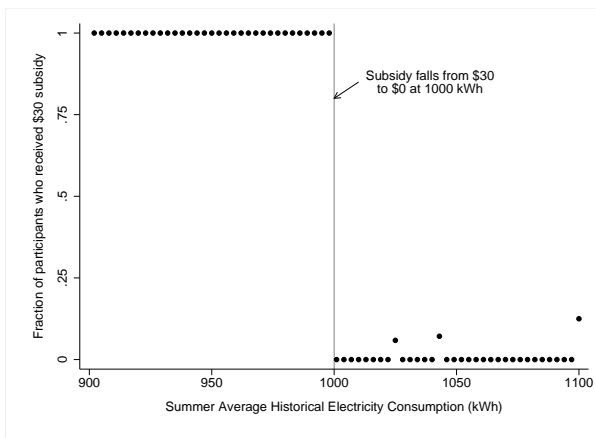
(a) Air Conditioners, 500 kWh Threshold



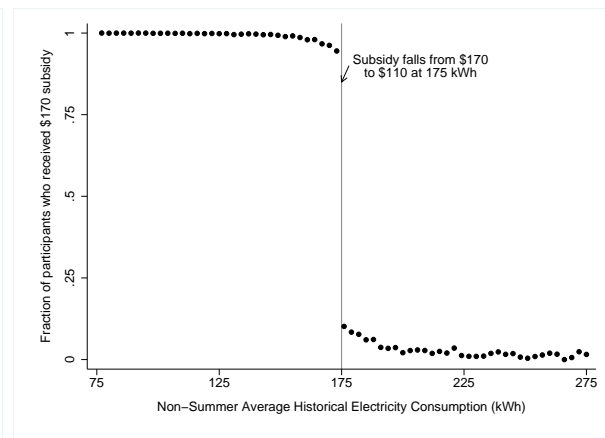
(b) Air Conditioners, 750 kWh Threshold



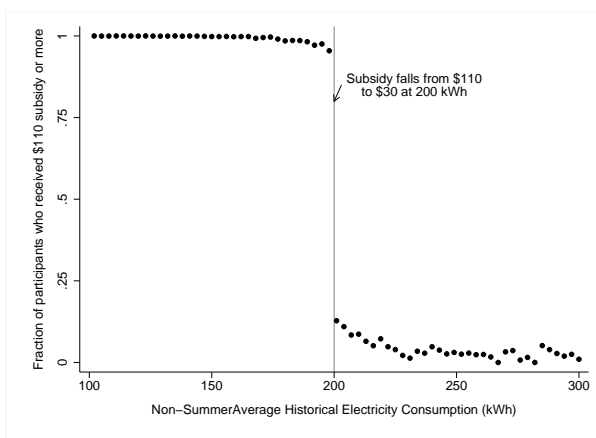
(c) Air Conditioners, 1000 kWh Threshold



(d) Refrigerators, 175 kWh Threshold



(e) Refrigerators, 200 kWh Threshold



(f) Refrigerators, 250 kWh Threshold

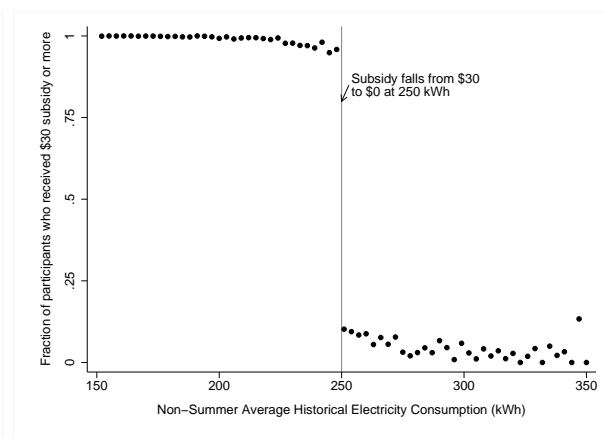


Figure 2.3: Smoothness of Running Variable Across Subsidy Thresholds

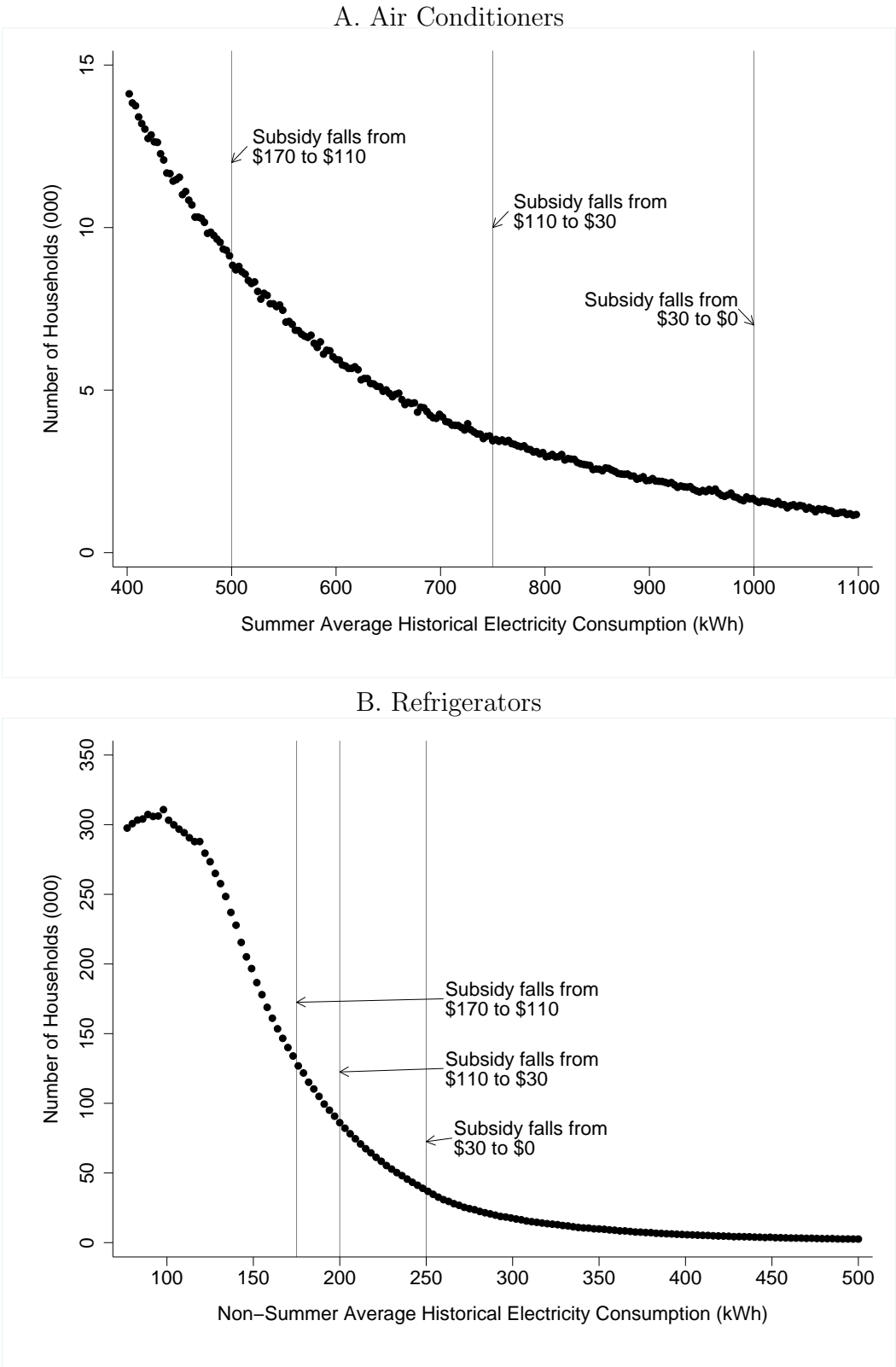


Figure 2.4: Smoothness of Household Income Across Subsidy Thresholds

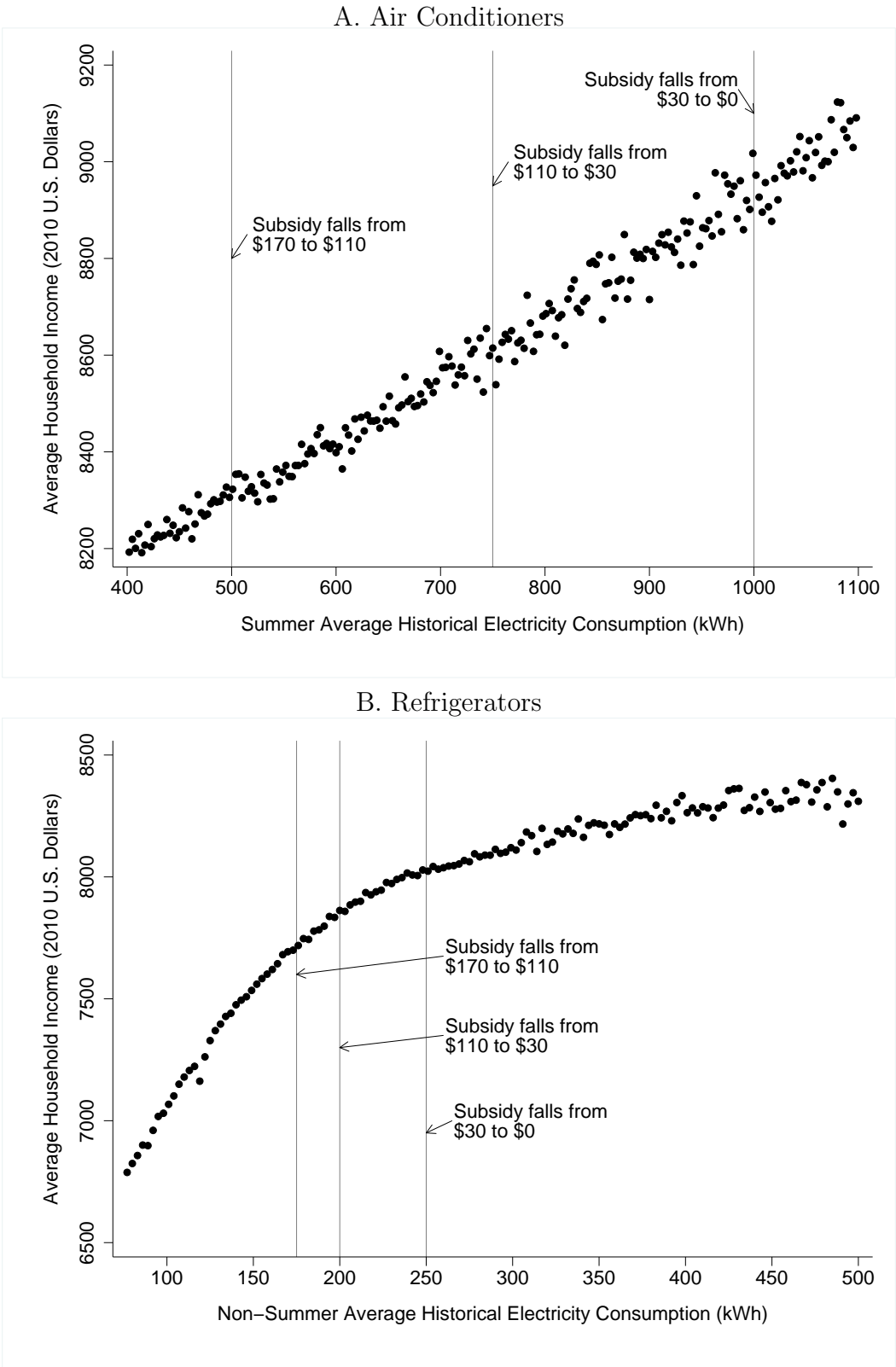


Figure 2.5: Program Participation



Table 2.1: Subsidy Amounts

Program Tier	Direct Cash Payments	Maximum Loan Amount	Consumption Levels, Air Conditioners	Consumption Levels, Refrigerators
1	2,200 Pesos (\$170)	3,400 Pesos (\$270)	251 – 500	76 – 175
2	1,400 Pesos (\$110)	4,200 Pesos (\$330)	501 – 750	176 – 200
3	400 Pesos (\$30)	5,200 Pesos (\$410)	751 – 1,000	201 – 250
4	No cash payment	8,700 Pesos (\$690)	1,000+	250+

Notes: This table describes the direct cash payments and maximum loan amounts available to households with different levels of average historical electricity consumption (in kilowatt-hours per month). For further details, see Appendix. Dollar amounts are reported in U.S. 2010 dollars using the average exchange rate for 2010 (12.645 Pesos per dollar). For expositional clarity, we rounded all dollar amounts to the nearest \$10. Households with consumption below the first tier were ineligible for subsidies.

Table 2.2: RD Estimates of the Effect of Subsidies on Program Participation

Subsidy Increase	Percent of Households Participating		Percent Change in Participation at the Threshold	Implied Slope	Implied Price Elasticity
	At Lower Subsidy Amount	At Higher Subsidy Amount			
(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Air Conditioners					
\$110 to \$170	1.45 (0.23)	1.75 (0.29)	20.6 (8.7)	0.0054 (0.0023)	0.88 (0.38)
\$30 to \$110	1.07 (0.23)	1.55 (0.31)	44.6 (11.6)	0.0069 (0.0019)	1.76 (0.49)
Panel B. Refrigerators					
\$110 to \$170	1.37 (0.11)	1.63 (0.13)	19.1 (2.7)	0.0047 (0.0007)	0.90 (0.13)
\$30 to \$110	0.89 (0.08)	1.20 (0.10)	34.1 (4.7)	0.0044 (0.0006)	1.51 (0.21)

Notes: This table reports sharp RD estimates of the effect of increased subsidies on program participation from four separate regressions. In each regression, the sample includes all households within our preferred bandwidth. We use a 100 kWh bandwidth for air conditioners, and a 50 kWh bandwidth for refrigerators. All regressions include a cubic polynomial in average historical electricity consumption, normalized to zero at the threshold. Column 2 reports the estimated intercept. Column 3 reports the estimated intercept plus the estimated coefficient on an indicator variable equal to one for households below the eligibility threshold. Column 4 reports the percent change between the previous two columns. Column 5 reports the change in the percent of households participating (as a fraction of all households at the threshold) per dollar of subsidy change. Column 6 reports the implied price elasticities evaluated using the net change in the cost of replacement at each threshold. Standard errors are clustered at the municipality level.



Table 2.3: Alternative Bandwidths and Specifications

Panel A: Air Conditioners					
Subsidy Increase	Sharp RD				Fuzzy RD
	Cubic Polynomial 125 kWh	Cubic Polynomial 100 kWh	Cubic Polynomial 75 kWh	Local Linear 50 kWh	Cubic Polynomial 100 kWh
\$110 to \$170	20.7 (8.4)	20.6 (8.7)	29.4 (9.8)	30.7 (8.9)	20.9 (9.1)
\$30 to \$110	56.1 (13.4)	44.6 (11.6)	38.2 (11.9)	49.4 (13.4)	45.7 (12.9)
Panel B: Refrigerators					
Subsidy Increase	Sharp RD				Fuzzy RD
	Cubic Polynomial 75 kWh	Cubic Polynomial 50 kWh	Cubic Polynomial 25 kWh	Local Linear 15 kWh	Cubic Polynomial 50 kWh
\$110 to \$170	22.5 (2.3)	19.1 (2.7)	15.7 (3.2)	15.0 (3.2)	22.6 (3.1)
\$30 to \$110	35.2 (4.5)	34.1 (4.7)	37.5 (6.2)	40.6 (6.1)	41.0 (5.7)

Notes: This table reports the estimated percent increase in program participation from 20 separate regressions, corresponding to the four main eligibility thresholds. The specification and bandwidth used are indicated at the top of each column. The cubic polynomial estimates using a 100 kWh bandwidth for air conditioners, and a 50 kWh bandwidth for refrigerators are identical to our estimates in Table 2.2. The columns on either side report estimates from two alternative bandwidths. The fourth column reports estimates using local linear regression with a uniform kernel. The final column reports estimates from a fuzzy RD specification which scales the estimated change in participation by the size of the discontinuity at the eligibility threshold. Standard errors in the first four columns are clustered at the municipality level and in the last column are block bootstrap by municipality with 5,000 repetitions. See text for details.

Table 2.4: Inferring the Fraction Non-Additional

	Projection Based on Linear Demand (1)	Projection Based on Elasticities (2)
Fraction Non-Additional	53.6% (4.8)	43.3% (6.0)
Average Payment Per Induced Replacement	\$328 (36.7)	\$269 (30.2)

Notes: In this table we use the RD estimates from the thresholds to infer what fraction of participants would have replaced their appliances with zero subsidy. For the average payment per induced replacement we divide total subsidy payments by the implied total number of additional participants. See text for details.

## Chapter 3

# The Effect of Biomass Electricity Generation on Forest Harvest in Maine

### 3.1 Introduction

Electricity produced by burning biomass, especially waste wood and farm residues, supplies a major share of renewable generation in the United States. Biomass provided 34% of non-hydroelectric renewable generation in 2010.<sup>1</sup> Much of the expected future growth in renewable electricity generation is also expected to come from biomass (Figure 3.1). These optimistic projections are driven by low costs relative to other renewable technologies, in large part because of the abundance of forestry and agriculture residues available to fire biomass generators.<sup>2</sup>

While biomass electricity is a promising source of low-cost renewable energy, its growth presents environmental challenges for forests. Demand for energy fuels may increase removals of low-value wood during harvests of other forest products. These increased wood removals often come through increased “whole-tree” harvesting, where the tops and branches of cut trees are collected instead of being left in the field. Demand for biomass fuel may also encourage the harvest of standing dead trees (which have low timber value) and “downed woody material” (DWM), which is woody material on the ground (Manomet Center for Conservation Sciences, 2010). If this additional harvest is not well-managed, it can damage forest health. The main concerns are reduced nutrient inputs to soils, increased erosion, and loss of habitat for sensitive species (Benjamin, 2010). In addition, if biomass generation encourages greater removals of standing trees, the medium-run carbon neutrality of biomass

---

<sup>1</sup>Energy Information Administration. 2011. Renewable Energy Consumption and Electricity Preliminary Statistics 2010. <http://www.eia.gov/renewable/annual/preliminary/>

<sup>2</sup>Energy Information Administration. 2012. EIA projects U.S. non-hydro renewable power generation increases, led by wind and biomass. <http://www.eia.gov/todayinenergy/detail.cfm?id=5170>

electricity may be in question (Manomet Center for Conservation Sciences, 2010).

This study's main contribution is to estimate an empirically-grounded price elasticity of supply for forest-derived biomass fuel in the US state of Maine. This parameter is central to understanding how increased demand for woody biomass fuel will affect forests. Figure 3.2 shows a simplified example of how increased demand for biomass fuels increases harvest, and how the amount of the increase demands on the slope of the supply curve. Estimates of the quantity of biomass removals due to energy demand can be combined with ecological models to understand the environmental impacts of biomass electricity generation on forests. The results of this study are also relevant for understanding the future potential of biomass electricity. The supply elasticity of wood fuels will affect the fuel prices faced by biomass power plants as the industry grows. These fuel prices are an important driver of the overall supply elasticity of biomass electricity.

I draw on an unusually detailed dataset that I compiled from annual forestry reports published by the Maine Division of Forestry. The data include 15 years of county-level quantity and prices for forest residues sold for biomass energy use. The availability of price data for biomass fuels is a distinguishing characteristic of this dataset that enables this analysis. To my knowledge, it is the most detailed publicly available data on biomass fuel prices in the United States.

I find that doubling the price of woody biomass in Maine increases harvest by about 64% on average. In the reduced form analysis, I also find that coal prices are a major determinant of woody biomass harvest, suggesting that environmental policies that raise the price of coal will affect forest health through increased biomass harvest. I do not find a statistically significant effect of subsidies to biomass power plants on biomass harvest.

## Biomass energy and forest harvest

Most of the fuel used in biomass generators in the United States comes from forests.<sup>3</sup> Woody biomass fuels are byproducts of forest harvest operations for higher-value sawlogs and pulpwood. Sawlogs are used to make boards, furniture, and other wood products. Pulpwood is made into pulp and paper at paper mills. Sawlog and pulpwood harvests generate otherwise unmarketable residues - for example, the tops and branches of cut trees.<sup>4</sup> Landowners and loggers decide between leaving these residues and other low-value woody materials (like standing dead trees), or collecting and selling them as biomass fuel. If they are to be sold, they are typically processed into wood chips and then trucked to power plants and industrial facilities with wood boilers.

The ecological impacts of woody biomass removals are complicated and site-specific. In some areas, including public lands in the Western US, increased removals may improve environmental outcomes by reducing wildfire risk. Removing low-value standing trees may

---

<sup>3</sup>Other fuel sources include farm residues, municipal waste, construction and demolition waste, and landfill natural gas

<sup>4</sup>In an informal interview, one logger suggested that the volume of wood in a typical cut in Maine is 50% pulpwood, 25% sawlogs, and 25% low-value residues.

also aid in stand management if debris or small trees slow forest regeneration. However, in some regions, particularly in the Eastern US, unregulated removals of low-value wood may harm the environment. Foliage and branches are a small share of a tree's volume, but can contain up to half of the stored nutrients, so that whole-tree harvesting may deplete soil nutrients (Benjamin, 2010). A number of threatened wildlife species are also thought to be impacted by removals of woody material and standing dead trees (Janowiak and Webster, 2010). Finally, if woody biomass harvest removes residues that would otherwise stabilize soils, water quality can be damaged by increased erosion (Janowiak and Webster, 2010).

These scientific findings have prompted regulatory responses. Massachusetts suspended support for biomass electricity in its renewable portfolio standard program in 2009, and is currently revising its wood harvest rules (North Carolina Clean Energy Technology Center, 2012). The European Union is also formally examining the environmental sustainability of woody biomass energy. Member states have proposed to rapidly increase biomass generation to meet an EU-wide target of 20% of electricity consumption from renewables by 2020 (European Commission Directorate General For Energy, 2011). There has also been vigorous debate in the Australian Parliament about whether to extend clean energy subsidies to generators that burn forest waste.<sup>5</sup> At the same time, forestry regulators in the US and elsewhere have begun to develop guidelines or regulations for woody biomass harvest (Evans and Perschel, 2009).

## The supply of woody biomass fuels

Green wood chips are bulky for the amount of energy they contain, making transportation expensive. In fact, transportation is often the largest share of the marginal cost of delivered biomass fuels. This means that markets for woody biomass fuels are local.

Most studies of markets for woody biomass fuels have used engineering simulation methods. In general, these studies conclude that there should be a large pool of readily available waste products from forestry and agriculture, making supply highly elastic with respect to price. The most prominent engineering study is the "Billion Ton Vision" from Oak Ridge National Laboratory (Perlack and Erbach, 2005; Department of Energy, 2011). This study finds that up to 555 million dry tons of biomass fuels are available nationally today, which is more than double what is currently used. For forestry residues, the supply curve is linear from about \$2 per dry ton to \$45 per dry ton<sup>6</sup>, with an implied US-wide price elasticity of roughly 4.75. Above \$45 per dry ton supply becomes inelastic, reflecting the constraints in their model (e.g., that 30% of residues must be left in the forest for environmental reasons). For Maine, the implied supply elasticity of forest residues is about 0.5 – 0.6.

These engineering studies depend on untested assumptions about the behavior of forest owners and farmers. This literature would benefit from empirical economic work using real

---

<sup>5</sup>Watt, J. 2012. Bid to use 'forestry waste' in biomass energy projects fails by one vote. Australian Broadcasting Corporation. March 20, 2012. <http://www.abc.net.au/local/stories/2012/03/20/3459509.htm>.

<sup>6</sup>Woody fuels are sometimes reported in dry tons and sometimes in green tons. An accepted approximate conversion factor is that green wood chips are 40-50% moisture by weight.

market data. To my knowledge, no econometric estimates of the supply of forest and farm residues for energy use exist. This is partly due to the scarcity of detailed price information.

While there have been few or no formal econometric studies of energy wood chip supply, there is a rich literature on the supply of other timber products. A major theme of many recent studies has been that the supply of timber on private lands is very price inelastic (Newman and Wear, 1993; Liao and Zhang, 2008). In contrast to the results found in engineering simulations, these results for other timber products have led some to question whether woody biomass production will respond significantly to increased prices (Manomet Center for Conservation Sciences, 2010).

### **Policies related to woody biomass electricity generation**

The market for woody biomass fuels is affected by two types of policies. The first is subsidies that affect consumers of woody biomass fuel - power plants and industrial facilities. The second is regulations governing suppliers of woody biomass fuels. This section discusses both types of policies in the US.

Biomass electricity generation has been supported by significant state and federal subsidies. Since beginning on January 1, 2005, the federal Producer Tax Credit (PTC) has paid 1.1 cents per kilowatt-hour (kWh) of electricity produced by “open-loop” biomass generators.<sup>7</sup> The subsidy is only for generation sold through the grid to an unrelated party, so industrial plants that generate their own electricity are not eligible.<sup>8</sup> It is illegal for industrial facilities to sell biomass power to collect the subsidy while simultaneously purchasing grid power to run the plant.

Several states also subsidize biomass electricity generation. The most common form of support is Renewable Portfolio Standards (RPS) that require utilities to procure a minimum fraction of their generation from renewables, including biomass. Several US states have RPS programs. In most programs, renewable generators are granted Renewable Energy Credits (RECs) in the amount of their generation. They can sell these RECs to utilities, who comply with the RPS by holding RECs equal to the required amount of renewable generation. RECs are traded in competitive markets, and prices have varied widely over time. In 2005, Massachusetts RECs were trading at over \$50 per megawatt-hour (MWh). In 2010 the market price was close to \$10/MWh (Wiser and Smith, 2007; Wiser, 2010).

On the supply side, biomass harvest on private lands is generally under the authority of state forestry regulators. Until recently, few states had addressed biomass harvest. Forestry regulations have principally been concerned with protecting water quality (e.g. protecting

---

<sup>7</sup>Closed-loop biomass electricity is made from dedicated energy crops grown specifically for electricity generation. Open-loop biomass is all other types of biomass electricity generation, including generation from wood waste.

<sup>8</sup>The subsidy program is described on IRS Form 8835: Renewable Electricity, Refined Coal, and Indian Coal Production Credit (2011). In practice, some industrial facilities do sell excess generation to the grid and collect the subsidy on those units. However, as I show below, the primary use of wood energy fuels at the industrial facilities in this study is to generate electricity for on-site use.

streams and limiting erosion). Many states, including Alabama, California, Maine, Michigan, Minnesota, Missouri, New Hampshire, Pennsylvania, and Wisconsin, have begun to develop voluntary best practices for sustainable biomass harvest. Most states suggest that a certain percentage (e.g. 20%) of tops and branches and standing dead trees be left (Evans and Perschel, 2009; Manomet Center for Conservation Sciences, 2010). For purely private costs like reduced nutrient input on private land, information provision may be sufficient to achieve an efficient outcome. However, if biomass harvest creates external costs by harming water quality, biodiversity, or public lands, then binding policy interventions may be welfare-increasing.

## The forest products and electricity industries in Maine

The forest products industry is a major component of Maine's economy. A large number of sawmills and paper mills produce a wide variety of wood products and employ thousands of people (Innovative Natural Resource Solutions, 2005). Maine also has nine biomass power plants with capacities ranging from 20 – 45 MW (Figure 3.3) and capacity factors in the range of 65% - 85%.<sup>9</sup> In addition to wood chips from forests, these plants may also burn construction debris and municipal woody waste. In general, however, they tend not to burn fuels other than wood. Several of these plants have been approved to sell RECs into the Massachusetts and Connecticut RPS markets.<sup>10</sup>

Besides power plants, the largest consumers of waste wood for energy are paper mills and sawmills. Maine has 12 large paper mills and dozens of sawmills as of 2010 (Liliehalm and Trospen, 2010). Unlike biomass power plants, these facilities consume the majority of the electricity they generate on-site, which makes them ineligible to receive the PTC subsidy on most of their generation. These plants can generally burn a wide variety of fuels and actively manage their fuel sources to minimize cost. A quote from Verso Paper about their paper mill in Bucksport exemplifies the flexibility with which paper mills substitute a wide variety of fuels:

“At Bucksport we use natural gas, oil, coal, biomass and tire-derived fuel (TDF)... We have the ability to adjust our fuel mix hourly in order to achieve the best result at the lowest cost.”<sup>11</sup>

Wood chips are also used in Maine to produce heat or electricity for large institutional facilities like schools. There is little published information about the prevalence of these

---

<sup>9</sup>A capacity factor is a measure of how much electricity a power plant produces relative to the maximum amount it could produce during that period.

<sup>10</sup>There is also a growing pellet industry in Maine that manufactures densified wood pellets for residential heating or for sale to large-scale electric generators in Europe, where carbon emissions regulations make wood an attractive generator fuel. These pellet mills use the same paper-grade pulpwood that goes to paper mills. The low-grade residues sold as wood chips are not suitable to their production process.

<sup>11</sup>Verso Paper Corporation. 2009. Maine on paper: An industry we can't afford to lose. <http://www.versopaper.com/OurCompany/Mills/TimeForAction/>

systems or about their ability to co-fire alternative fuels such as coal. The “Governor’s Wood-to-Energy Task Force Report” (State of Maine, 2008) describes a number of applications of wood chip heating in commercial and institutional facilities.

The environmental impacts of harvesting woody biomass fuel for these facilities are similar to the rest of the US. Benjamin, Lilieholm, and Coup (2010) find that increased use of biomass in Maine merits more careful regulation, especially wildlife protection. In a small study of 12 Maine sites, Briedis et al. (2011) find that whole-tree harvests in 2007-2008 complied with best practice guidelines for the quantity of residuals left in the field. However, the amount of large-diameter woody material and standing dead trees were consistently below best practice recommendations for wildlife protection.<sup>12</sup>

The state of Maine has also studied woody biomass harvesting. The Maine Forest Service commissioned a consultant report that used engineering models to show that biomass harvests could increase substantially while still leaving some material in the field (Maine Forest Service, 2008). The state issued voluntary, qualitative guidelines for woody biomass harvest in 2010 (Benjamin, 2010).

## 3.2 Model and instruments

To motivate possible instruments for price in the supply equation for wood chips, this section presents a highly simplified model of demand and supply for woody biomass fuels in Maine. I discuss demand and supply in this market, propose two candidate instruments, and then present a proposed regression model.

### Demand for woody biomass at power plants, paper mills, and sawmills

Wood chips are an input to production in power plants, paper mills, sawmills, and other facilities with wood-fired boilers. I assume that these firms choose output to maximize profit from their facilities:

$$\pi_d = (p^g + s)Q^g - C_f(p^w, p^{subst}, Q^g) - C_r(Q^g, p^p, p^l) \quad (3.1)$$

$Q^g$  is the facility output (electricity for power plants, paper for paper mills, etc).  $p^g$  is the exogenous per-unit price of this output.  $s$  is the per-unit output subsidy (for example, subsidies for biomass electricity generation on a per-kWh basis).  $C_f$  is the cost of fuel used by the facility. This is a function of output and the prices of the fuels which the facility can use. For any given facility, this is some subset of wood chips ( $p^w$ ) and substitute fuels, including coal and oil. I assume that the facility chooses the fuels compatible with its generating equipment that minimize its cost of output. I do not address long-run modifications to

---

<sup>12</sup>The study is unable to attribute this deficiency to recent biomass harvests vs. previous harvesting activities at these sites.



generating equipment to burn different fuels. I also assume that the firm is a price taker in the output and wood chip markets.  $C_r(\cdot)$  captures all other costs besides wood fuel. For paper mills and sawmills, this includes the cost of pulpwood or sawlog inputs that are made into paper and wood products. These prices are  $p^p$  and  $p^l$ , respectively.

Under this extremely simple model, the firm maximizes profit by producing where marginal cost equals the market price of the output plus the per-unit subsidy. This suggests that demand for wood chip fuel should be a function of output prices for paper, boards, and electricity; of the level of subsidies to biomass generation; of the prices of pulpwood and sawtimber; and of the prices of wood chips, coal, and oil.

### Supply of woody biomass by forest owners

I assume that forest owners maximize profit from production of sawlogs, pulpwood, and wood chips. Wood chips are a byproduct of harvesting the first two products. After producing any amount of sawlogs and pulpwood, the owner chooses how much otherwise unmarketable leftover wood to chip and sell for energy. For simplicity, I assume that forestry firms are price takers for wood chips. I abstract away from the dynamic optimization problem of optimal timber rotation (e.g., the Faustmann model) because my main interest is in motivating instruments for the wood chip byproduct, not in describing the supply of timber.

The profit function for wood chip production is:

$$\pi_s = Q^w p^w - C(Q^w, T(p^{oil}, d), Q^l, Q^p) \quad (3.2)$$

As above,  $l$  indexes sawlogs,  $p$  indexes pulpwood, and  $w$  indexes wood chips. The cost of wood chip production  $C(\cdot)$  depends on the amount of wood chips produced, the amount of primary products (logs and pulpwood) produced, and transportation costs  $T$ .  $T$  is a function of fuel costs, which vary in time with the price of oil  $p^{oil}$  and distance  $d$  from the harvest site to the biomass energy user. This distance does not vary in time.<sup>13</sup>

This very simple model suggests that wood chip supply should be increasing in wood chip prices; increasing in pulpwood and sawlog quantity; and decreasing in gasoline prices, labor costs, and other harvest costs.

### Market clearing

The final equation to complete the model is the market clearing condition:

$$Q_{demand} = Q_{supply} \quad (3.3)$$

### Candidate instruments

I am interested in the effect of  $p^w$  on  $Q^w$  in the supply equation (equation (3.2)). As described by Working (1927), price and quantity are endogenous but the true effect can be recovered

<sup>13</sup>As I show below, at least for power plants, capacity is fixed during the study period

by instrumenting for price with demand shifters. The simple model above suggests several possible instruments.

One possible instrument is subsidies to biomass electricity generation. In my simple model, subsidies will be positively correlated with woody biomass prices. Assuming plants are not already operating at full capacity, output subsidies will increase electricity production, increasing their demand for wood chip fuel. The exclusion restriction is also clear. Because the subsidy is introduced exogenously and goes directly to power plants for their output of electricity, it should not be correlated with any unobserved determinants of wood chip supply.

The subsidy program with the most promise as an instrument is the federal Producer Tax Credit. Unlike RECs, which require firms to self-select into “treatment” (the opportunity to receive subsidies), the PTC applies automatically to all biomass facilities that sell electricity over the grid. In addition, the level of the subsidy is known, unlike in REC markets where prices are usually not available.

Because some counties contain paper mills and sawmills but no biomass power plants, and transportation costs for wood chips are high, there is quasi-experimental variation in eligibility for the Producer Tax Credit subsidy between local wood chip markets. Table 3.1 confirms that power plants sell a much larger share of their generation to the grid than industrial plants. This cross-sectional variation is another reason that the PTC is an attractive instrument.

A second promising set of instruments is the prices of substitute fuels for electricity generation. Paper mills, sawmills, and other facilities that burn wood for energy typically have boilers that can burn multiple fuels. These mills choose their fuels based on current prices and availability (Pinkerton, 2007). This means that the price of wood chips should be positively correlated with the prices of these other fuels.

The exclusion restriction for substitute fuels requires that fuel prices do not affect wood chip supply by changing input demand for pulpwood and sawtimber. A sufficient condition would be that fuel prices only affect choices between fuels, not the overall output of the facility. There is strong support for this assumption. Output from Maine’s paper mills was relatively flat during the study period, with the exception of a decrease in 2009. Industry leaders generally agree that paper and log output is exogenously determined by global demand (Price Waterhouse Coopers, 2011). Even if energy prices in Maine did depress mill operations in some years, it is typical to ship pulpwood and sawlogs long distances to markets, likely reducing any impact of local processor demand on local pulpwood and sawlog quantities. For all of these reasons, I argue that there is a valid exclusion restriction and that substitute fuel prices are a good instrument.

To decide which substitute fuels would make good instruments, I used data from the US Energy Information Administration Generator Operations Reports from 2001 to 2010 (as far back as the data are available in a consistent format) to identify all of the fuels used at power plants and paper mills in Maine larger than 1 MW that burn wood.<sup>14</sup> As Figure 3.4 shows,

---

<sup>14</sup>Energy Information Administration. 2001 – 2010. Form 923: Generator Operations Reports [10 separate

the primary alternatives to wood chips at these facilities during the study period were black liquor, coal, and oil.<sup>15</sup> Black liquor is spent liquids from the pulping process at paper mills. It is used by the plants that create it, so there is no price data to use as an instrument. Oil price is not a good instrument either because it enters the supply equation through gasoline costs. The price of coal, however, is a good candidate. It is traded in a national market, so its price can plausibly be considered exogenous. Table C.1 in the Appendix shows that coal prices are uncorrelated with two other possible determinants of wood chip supply that are not included in my regressions.

A third category of instrument suggested by equations (3.1) and (3.2) is the wholesale price of electricity. However, biomass power plants often sign long-term, confidential power purchase agreements with electric utilities, making it difficult to observe the output price faced by the plant. I avoid instrumenting with electricity prices for that reason.

## Proposed Regressions

First, I propose the following reduced form model where I regress the outcome variable on the instruments and other exogenous covariates:

$$\ln(Q_{it}^w) = \beta_0 + \beta_1 \ln(p_{it}^{coal}) + \beta_2 PTC_i * after_t + \beta_3 after_t + \beta_4 \ln(p_{it}^{oil}) + \beta_5 \mathbf{time} + d_i + \epsilon_{it} \quad (3.4)$$

In this regression,  $i$  indexes counties and  $t$  indexes year.  $Q^w$  is annual wood chip quantity per unit of land area in the county.  $p^w$  is wood chip price.  $p^{coal}$  and  $p^{oil}$  are coal and oil prices.  $PTC$  is an indicator for whether the county contains a biomass power plant eligible for the Producer Tax Credit.  $After$  is an indicator for years after the subsidy took effect.  $d_i$  is a set of county dummies.  $\epsilon_{it}$  contains an unobserved year-specific shock across counties and unobserved idiosyncratic variation at the county-by-year level.<sup>16</sup>

**Time** is a vector of time trends estimated separately for treatment and control counties. In my preferred specification I include the year of the panel interacted with an ‘ever-treated’ indicator (so,  $time * PTC$  and  $time * control$ ). Because the dependent variable is in logs, this estimates an exponential time trend for each group.

My discussions with industry participants indicate that biomass plants source most of their wood from within 25 miles of the plant. In my main specification, I consider counties within 25 miles of a biomass power plants to be treatment counties for the Producer Tax Credit and the remaining counties to be control counties. This distance is an important

---

annual reports]. <http://www.eia.gov/>

<sup>15</sup>One facility used a large amount of natural gas, but I dropped it from the analysis because it was the only facility using natural gas and because its wood consumption was very small relative to its gas consumption.

<sup>16</sup>In the appendix, I also include annual sawlog and pulpwood production as covariates. I do not include them in my main specification because of endogeneity concerns: the prices and quantities of pulpwood and sawlogs processed at industrial facilities are determined in the same market as the prices and quantities of the wood chips that those facilities burn for energy.

identifying assumption for the difference-in-differences model. In Table C.5 I show how the results for the effect of the PTC change with different distance assumptions.

To estimate the price elasticity of wood chip supply, I use two stage least squares with the price of coal and the Producer Tax Credit as instruments. The first stage is

$$\ln(p_{it}^w) = \gamma_0 + \gamma_1 \ln(p_{it}^{coal}) + \gamma_2 PTC_i * after_t + \gamma_3 after_t + \gamma_4 \ln(p_{it}^{oil}) + \gamma_5 \mathbf{time} + d_i + u_{it} \quad (3.5)$$

and the second stage is

$$\ln(Q_{it}^w) = \alpha_0 + \alpha_1 \ln(\hat{p}_{it}^w) + \alpha_2 \ln(p_{it}^{oil}) + \alpha_3 after_t + \alpha_4 \mathbf{time} + d_i + v_{it} \quad (3.6)$$

This model is identified by time-series variation in the price of coal and by cross-sectional and time-series variation in output subsidies to power plants.

Because the price of coal does not vary between counties, there is a good reason to cluster my standard errors by year to allow for arbitrary patterns of correlation within years. At the same time, there is a strong rationale for clustering by county to allow for arbitrary patterns of serial correlation within counties. I cluster at the county level because this approach generates the larger standard errors and is thus more conservative. In future drafts I plan to explore the method of two-way clustering proposed by Cameron, Gelbach, and Miller (2011).

### 3.3 Data

#### Data on wood chips and other forest products

This study uses panel data on forest products prices and quantities for Maine’s 16 counties from 1996 to 2010. This is an unusual level of detail for forestry residues, which makes this an attractive dataset for understanding supplier behavior in these markets. To my knowledge, it is the only publicly-available dataset that includes detailed information on woody biomass prices across time and space. I compiled these data from the annual Wood Processor Reports and Stumpage Price Reports published by the Maine Forest Service.<sup>17</sup> The state gathers these data from the major wood producers and processors in each county. At the end of each calendar year, all major wood processors, including paper mills, sawmills, and biomass power plants, are required to submit a record of the wood that they have used by county of origin. Landowners are required to report the prices at which they have sold dozens of species and varieties of wood products. The state then aggregates this information in these annual reports.

The data include price and quantity for several woody biomass fuels, including biomass chips, “hog fuel” (mill residues), wood pellets, and construction and demolition waste. Biomass chips represent 90% of the production and are the only product for which prices are reported. The definition of biomass chips given in the annual reports is chips “produced in

---

<sup>17</sup>The individual annual reports are available from the Maine Forest Service at <http://www.state.me.us/doc/mfs/pubs/annpubs.htm>.

the woods using entire tree, stem or bole, branches and tops. These chips usually go to wood to energy facilities; however they are suitable for sludge composting, play ground padding, and mulch.”

The price data are stumpage prices, which are the prices received by landowners for harvested timber. For wood chips, stumpage prices are typically well below delivered prices. The mean price value observed in my data is \$1.90/ton, while delivered prices may be closer to \$15/ton. This is due to the costs of hauling biomass to the road, chipping it, and trucking it to wood chip consumers. Ideally this study would also use data on delivered prices in order to understand the contracting dynamics between landowners, logging firms, and final customers of wood chips. In the absence of such data, I assume that landowners sell wood residues into competitive markets.

The 15 years of panel data for 16 counties total to 240 observations. They are complete for quantity. There are 35 missing values for price in county-year observations where three or fewer firms reported wood chip transactions (these are not reported for confidentiality reasons). Because these are likely to represent lower-demand observations, I replace the missing values with the minimum price observed in that year for the corresponding PTC treatment category (power plant or no power plant). In Table C.3, I show that my results are robust to different ways of dealing with missing values, including dropping all observations without price data. I drop one outlier observation in 2004 with a reported price more than double the price reported in any other county that year. This leaves me with 239 total observations.

Figure 3.5 plots the wood chip production data. The left axis describes mean wood chip output per square mile in counties with and without biomass power plants. The dotted vertical line represents the first year of the Producer Tax Credit to biomass plants. As I will discuss below, the dashed red line corresponding to the right axis is the annual average coal price.

I do not observe wood chip storage. However, storage is unlikely to affect my annual production estimates. Biomass generators typically receive multiple fuel deliveries every day. Most biomass power plants store 15 - 60 days worth of fuel supplies (Wiltsee, 2000). Storage is constrained by the amount of land required for the wood chip pile, and by the fact that wood chips begin to decay and lose energy value after about two months. In addition, plants have an incentive to avoid depleting their reserves to buffer against unexpected supply disruptions.

One important question for the PTC portion of this study is whether the counties without biomass power plants are good “controls” for the counties that do have them. Table 3.2 compares power plant and control counties. As shown in the top panel, both types of counties contain facilities that use woody biomass for energy (paper mills, sawmills, and pellet mills). The second panel compares price and quantity of forest products in these counties before the introduction of the subsidy. The price and production of wood chips is not significantly different between power plant counties and controls during this period. However, the levels of production of other forest products are higher in power plant counties.

## Data on power plant locations and capacity

Data on the locations and capacities of biomass power plants eligible for the Producer Tax Credit subsidy come from the US Energy Information Administration (EIA). Biomass power plants for this study were considered to be plants whose primary fuel was “wood waste and wood chips” (“WDS” in the EIA fuel system) and whose primary purpose was electricity generation according to the North American Industry Classification System. This definition excludes plants whose primary purpose was generation for on-site industrial use.

I used EIA form 860 data to locate biomass plants meeting this definition. I found 11 total plants. To investigate whether capacity was fixed over the panel, I examined form 860 data for all available years (2001 - 2009). Nine of the 11 plants are continually present with the same nameplate capacity in the data, and all of them began operation prior to 1996, the first year of the panel. The other two plants were retired prior to the introduction of the Producer Tax Credit (in 2001 and 2003). The list of biomass power plants selected using this procedure matches the list of “stand-alone biomass-fired electrical generating facilities” provided in the official Directory of Maine Wood Processors (Lilieholm and Trospen, 2010).

## Data on substitute fuel quantities and prices

The data on fuel use at individual facilities are from the US Energy Information Administration (EIA). I compiled annual data from EIA Form 923, the “Power Plant Operations Report”. This report is filed by all facilities with greater than 1 MW of generation capacity. To create Figure 3.4, I used data for all facilities that used “wood and wood waste solids.” Data on coal prices are also from EIA. I used the “Total” price history for coal from the 2011 Annual Energy Review. This is the average annual price of U.S. coal at the point of first sale, excluding transportation costs (also called the free-on-board price). Figure 3.5 shows the coal price over time.

## 3.4 Results

This section presents my main results. I first show the results of a formal test of parallel trends in the pre-PTC period to support the identification for the PTC instrument. I then show a reduced form regression using both proposed instruments. Finally, I proceed to estimate the price elasticity of wood chip supply via two stage least squares. I present estimates using both instruments and also using coal price as the only instrument, since I find that the PTC instrument is relatively weak.

### Specification Test

Because the PTC instrument relies on a difference-in-difference identification strategy, I perform a formal test of parallel trends in the pre-treatment period. Table C.2 in the Appendix shows a regression of wood chip quantity on year\*treatment dummies, year dummies,

and county dummies for the years 1996 - 2004. None of the year\*treatment dummies are statistically significant.

### Reduced form regressions

Table 3.3 reports the regression of the log of wood chip quantity on the exogenous variables, exponential time trends, and county fixed effects. I find that the price of coal has a large effect on wood chip harvest. This effect is statistically significant at the 10% level. I do not find a statistically significant effect of the Producer Tax Credit subsidy on the quantity of wood chips sold. As I explore in the Robustness Checks section below, the estimated effect of the PTC depends somewhat on assumptions about the distance from which processors source wood.

### First stage regressions

Table 3.4 reports the regression of the log of wood chip price on the instruments, covariates, exponential time trends, and county fixed effects. The price of coal has a very strong first stage, and the PTC is weakly significant (at the 10% level). The joint F statistic for the two instruments is 72.

Coal prices have a very large effect on wood chip prices. As discussed previously, generators in industrial and commercial facilities trade off between coal and wood chips based on price. Additionally, since the marginal costs of coal and wood chip-fired generators in this region are very similar (Gan and Smith, 2006; Northeast Energy Solutions, 2008), rising coal prices could cause some coal and biomass generators to swap places in the industry supply curve. Figure C.1 in the Appendix shows suggestive evidence that generation from some biomass plants increased as coal prices rose.

I also find that the PTC may have had a small positive effect on the price of wood chips, but this result is only statistically significant at the 10% level once I include county fixed effects. My noisy results for the PTC subsidy in both the reduced form and the first stage could be explained in several ways. One possibility is that power plants source wood from greater distances than I have assumed. There are anecdotal reports of power plants purchasing woody biomass fuel from up to 100 miles away in some cases. At this distance, all counties in Maine would be affected by the policy, including the ones that I have considered control counties in this analysis.

### Two stage least squares regressions

Table 3.5 shows my price elasticity estimates. Because the price of coal seems to be a much stronger instrument than the PTC subsidy, I report results from using both instruments and from using only coal. I also report the OLS estimate for comparison.

OLS underestimates the price elasticity relative to the IV models. Using both instruments, the estimated elasticity is about 0.43. The advantage of using both instruments is

that the PTC includes cross-sectional as well as time-series variation. However, coal prices have much stronger first stage and reduced form effects than the PTC, so that my preferred specification is the “IV, coal only” model. This gives a supply elasticity of about 0.64. This estimate is similar to the engineering estimates for Maine from the Billion Ton Report (Department of Energy, 2011).

## Robustness checks

This section explores the sensitivity of these main results to different ways of handling missing price data, alternative time trends, and different assumptions about the distance from which power plants source wood chips. In general, the results are robust to changing assumptions.

Table C.3 in the Appendix shows that the results are stable with changes in how I handle missing data. Dropping 35 observations with missing price observations decreases the precision of my first-stage estimate for the PTC.

Table C.4 in the Appendix explores an alternative time trend specification. My main results include exponential time trends for treatment and control counties. Here, I do the regressions in levels and include a linear time trend. The mean price elasticity implied by the 2SLS model is about 0.67, which is close to the result from my main specification.

Table C.5 in the Appendix shows how changing assumptions about the distances from which power plants purchase wood affects my the reduced form estimate of the effect of the Producer Tax Credit subsidy on the quantity of wood chips produced. I try distances up to 40 miles (at greater distances, essentially all counties become treatment counties). In no case do I find a statistically significant effect of the program on the quantity of wood chips sold.

Table C.6 in the Appendix shows how the results change when I include pulpwood and sawlog production as covariates. Adding sawlog production does not change my elasticity results; adding pulpwood production decreases the price elasticity estimate slightly. In the reduced form, conditioning on pulpwood and sawlog production does not increase the precision of my estimates enough to draw conclusions about the effect of the PTC (analysis not shown).

## 3.5 Conclusion

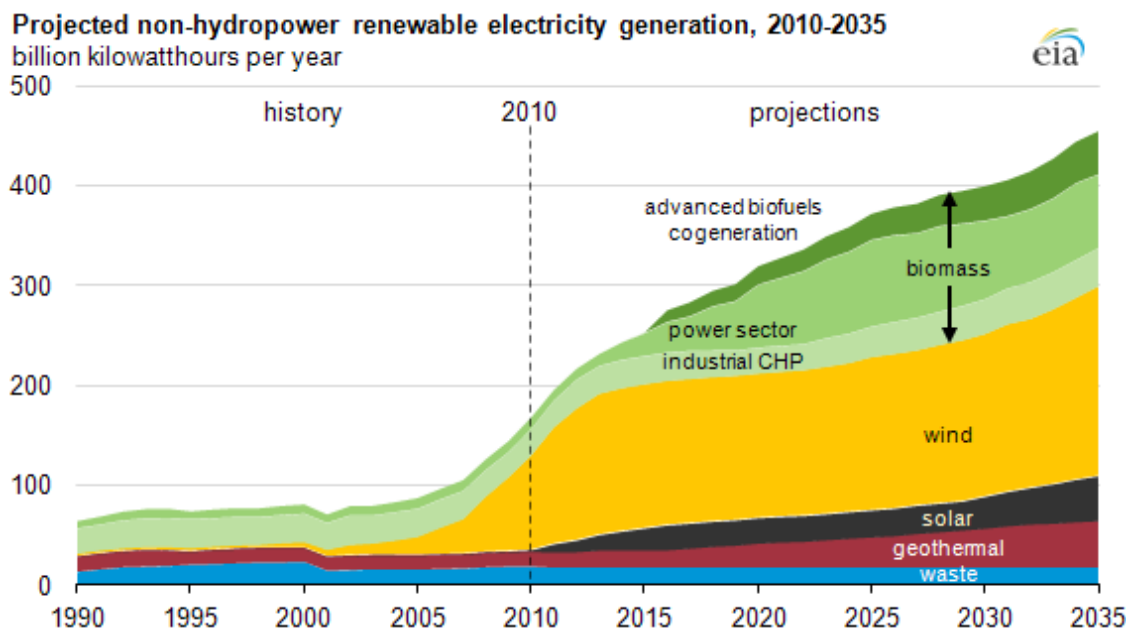
I find that the price elasticity of supply for woody biomass is about 0.64. On average, doubling the price of woody biomass increases harvest by about 64%. This is similar to engineering estimates (Department of Energy, 2011). This elasticity estimate can be used in a number of interesting and important ways. However, it is important to note that this estimate applies to the range of prices and quantities observed during this study. Extrapolating far from this range could lead to incorrect projections. The true marginal cost of wood chips is likely non-linear, with a rapid increase once available waste products are exhausted.



Combined with estimates of demand elasticities, the results of this study can be used to understand how shifts in demand for woody biomass affect the quantity and price of woody biomass harvested. Geographically-specific ecological models can be used to describe the environmental impacts of the predicted harvest quantity. Similarly, policymakers and power plant owners can use these estimates to better understand their future fuel costs. This combination of economic and scientific analyses should be an important part of the policy discussion around expanding biomass generation capacity in forest regions.

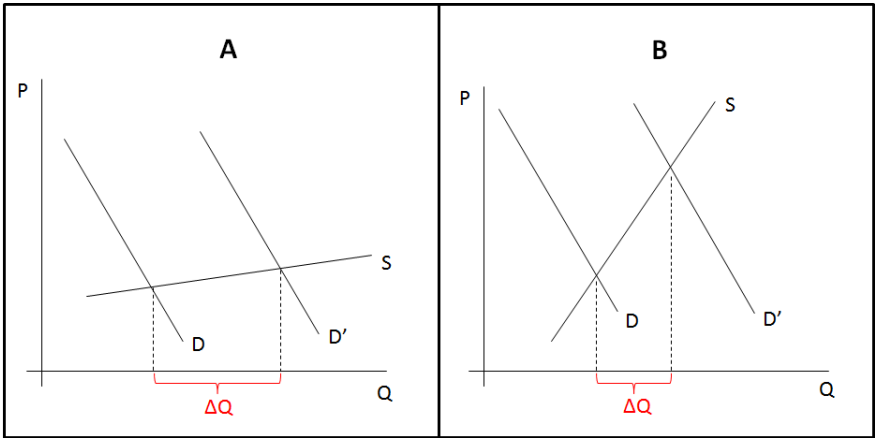
I use an instrumental variables strategy to measure the biomass supply elasticity. The instruments themselves are also of independent interest. I find that the price of coal is an important determinant of woody biomass harvest. On average, doubling the price of coal causes the quantity of woody biomass harvested to rise by 110%. The mechanism for this effect is either fuel-switching by commercial and industrial plants or increased competitiveness of biomass power plants in the electricity market. This finding is interesting in itself because it suggests that taxes on coal (for example, a carbon tax) would indirectly affect forests. I also find a small positive impact of the Producer Tax Credit subsidy, but these estimates are very imprecise.

Figure 3.1: EIA projections showing growth in biomass electricity



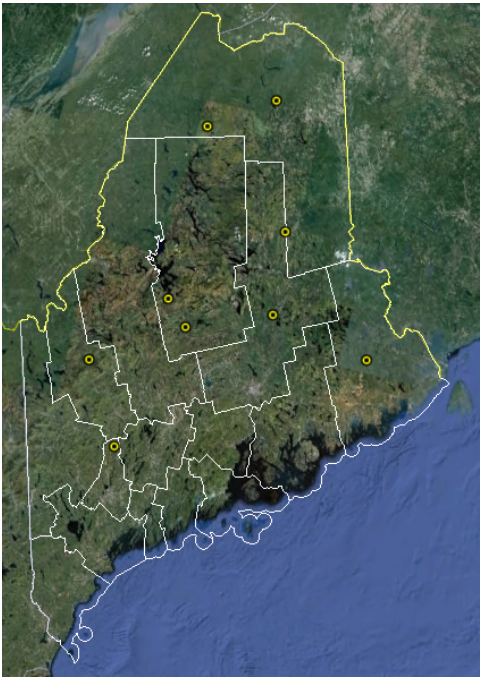
This figure is taken from: Energy Information Administration. 2012. EIA projects U.S. non-hydro renewable power generation increases, led by wind and biomass. <http://www.eia.gov/todayinenergy/detail.cfm?id=5170>

Figure 3.2: Effect of a parallel shift in woody biomass demand



This figure shows the effect of a parallel shift in woody biomass demand under different supply elasticities (for example, in response to a subsidy for renewable energy generation). In Panel A, supply is relatively elastic (line S). A demand shifts from D to D' results in a relatively large change in woody biomass harvest. In Panel B, supply is relatively inelastic. The same demand shift causes a smaller change in woody biomass harvest. Both panels take the slope of the demand curve as given. As explained in the text, demand for woody biomass fuel should be expected to be price elastic.

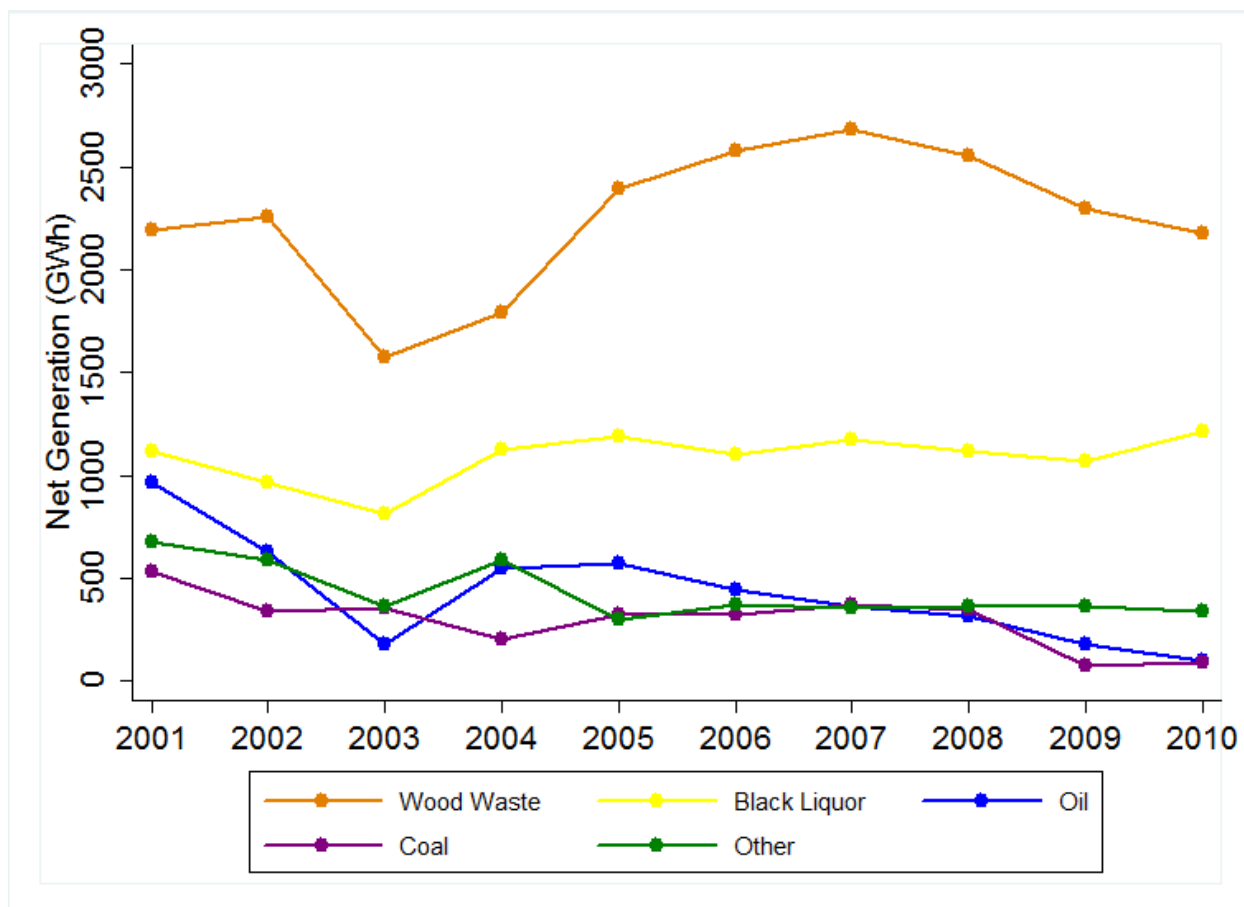
Figure 3.3: Map of biomass plants in Maine



---

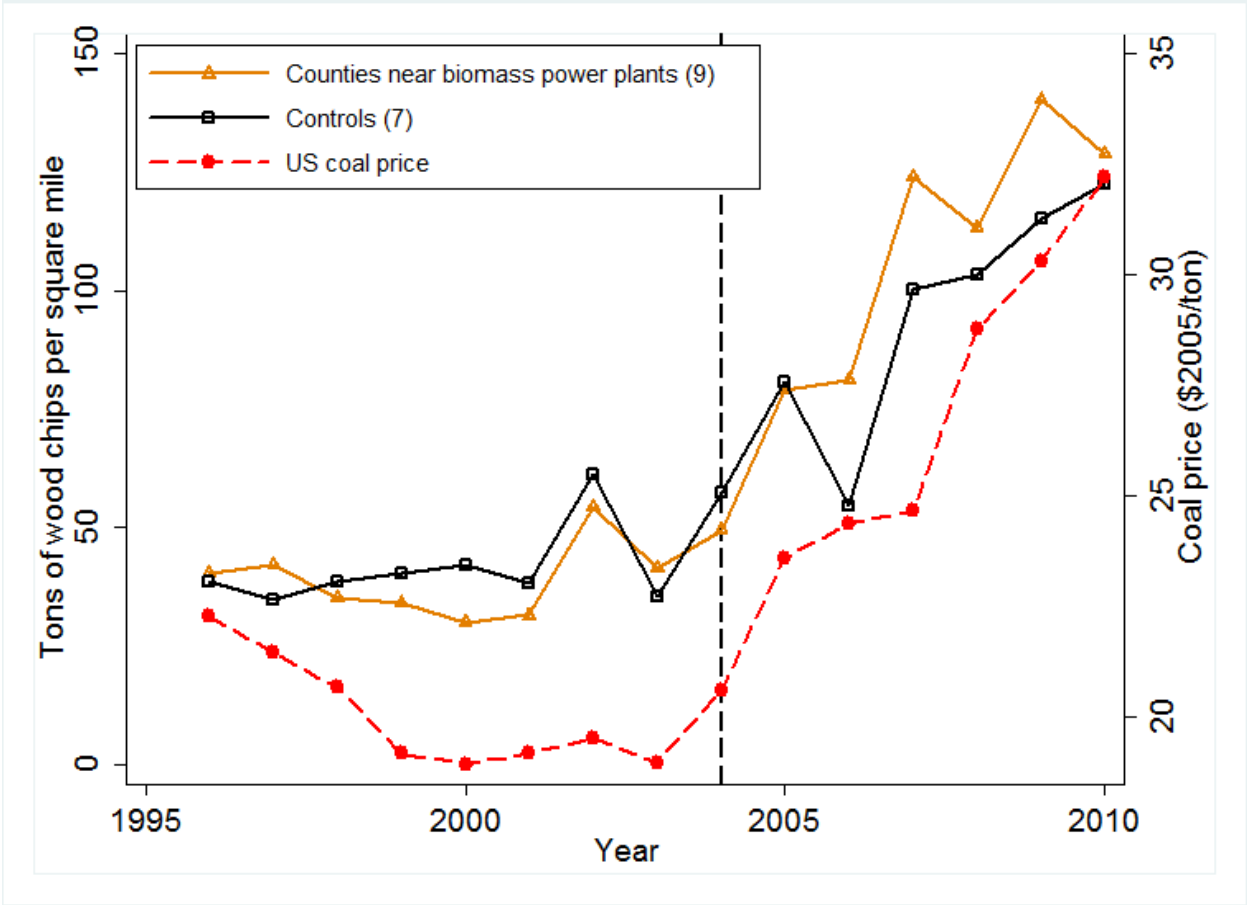
This map shows the locations of the nine active biomass power plants in Maine, all of which have capacities between 20 and 50 MW. White lines indicate county boundaries.

Figure 3.4: Primary fuels at Maine facilities that burn wood



This figure shows the net generation from different fuels at power plants, paper mills, and other large industrial facilities in Maine that burn at least some wood waste. The data are from EIA Form 923: Power Plant Operations Report. This excludes Verso Paper, which was the only facility to burn natural gas (and which had a very small usage of wood relative to its gas usage). ‘Other’ includes municipal waste, tire-derived fuels, purchased steam, and landfill gas.

Figure 3.5: Wood chip production by year



The left axis of this figure shows annual average wood chip production for counties within 25 miles of biomass power plants vs. other counties in Maine. The right axis shows the annual average price of coal. The vertical line represents the last year before introduction of the federal Producer Tax Credit for biomass power plants on Jan. 1, 2005. Wood chip data were compiled from annual Wood Processor Reports from the Maine Forest Service. Coal price data are the “Total” coal series from the US EIA Annual Energy Review.

Table 3.1: PTC eligibility at power plants and industrial facilities

	Power plants		Industrial facilities		p-value
	Mean (SD)	N	Mean (SD)	N	
Share of generation sold to the grid	0.90 (0.02)	28	0.20 (0.28)	35	0.00

This table shows the average share of electricity generation sold to a third party from wood-burning power plants and industrial facilities in Maine. This is calculated as “resale” divided by “total sources.” The data are available for 2007 - 2010 from EIA Form 923.

Table 3.2: Comparing counties near biomass power plants with control counties

	Treatment		Control		p-value
	Mean (SD)	N	Mean (SD)	N	
<i>Non-power plant users of wood energy</i>					
Number of Sawmills in 2010	16.3 (7.0)	9	11.3 (5.9)	7	0.15
Number of Paper Mills in 2010	1.1 (1.5)	9	0.3 (0.5)	7	0.20
Number of Pellet Mills in 2010	0.4 (0.17)	9	0 (0.0)	7	0.04
<i>Pre-treatment forest products output</i>					
Wood Chip Price (\$/green ton)	1.49 (0.73)	80	1.28 (0.55)	63	0.07
Wood Chips ('000 tons/sq. mile)	39.6 (27.8)	80	42.9 (30.7)	63	0.50
Pulpwood ('000 tons/sq. mile)	105.4 (37.7)	80	70.8 (25.3)	63	0.00
Sawtimber ('000 tons/sq. mile)	31.5 (11.1)	80	21.7 (10.4)	63	0.00
<i>Demographic information (2010)</i>					
Population density (per sq. mile)	56.3 (73.4)	9	135.3 (97.6)	7	0.10
Median HH income (\$000)	39.5 (4.0)	9	49.7 (5.3)	7	0.01

This table compares counties near biomass power plants ('treatment') to the other counties in Maine ('control'). The top panel reports the average number of industrial facilities that could use wood chips for energy in treatment and control counties. These numbers are for 2010 (the only year for which comprehensive statistics are available), and come from the *Directory of Maine's Primary Wood Processors* (Liliehalm and Trospen, 2010). The middle panel show forest products statistics from the years before the federal Producer Tax Credit for biomass power plants. The bottom panel includes demographic information from the 2010 US Census. 'p-value' is the the p-value of a t-test of similarity of means.



Table 3.3: Reduced Form: Effect of coal price and the PTC on wood chip quantity

Dependent variable: $\ln(\text{wood chips per square mile})$						
$\ln(\text{Coal Price})$	1.16	(0.61)*	1.16	(0.61)*	1.15	(0.60)*
PTC*after	0.03	(0.18)	0.04	(0.21)	0.04	(0.21)
PTC	-0.04	(0.21)				
after	0.21	(0.22)	0.22	(0.26)	0.25	(0.28)
$\ln(\text{Oil Price})$					-0.06	(0.12)
Constant	-0.27	(1.74)	-0.29	(1.70)	-0.08	(1.65)
Exponential time trends	Yes		Yes		Yes	
County Fixed Effects	No		Yes		Yes	
Counties			16		16	
Observations	239		239		239	

This table reports OLS regressions of the log of wood chip quantity on the instruments and covariates. Standard errors are clustered at the county level. The data are a 15-year panel created from Maine Forest Service annual reports. ‘Coal Price’ is the annual average price of coal. ‘PTC\*after’ is an indicator for counties near biomass power plants in years the Producer Tax Credit was available. ‘PTC’ is an indicator for counties near biomass power plants. ‘After’ is a dummy for years after introduction of the PTC. All regressions include separate time trends for counties with and without the PTC. Legend for significance stars: \* 10%; \*\* 5%; \*\*\* 1%.

Table 3.4: First stage: Effect of coal prices and the PTC on wood chip prices

Dependent variable: ln(wood chip price)						
ln(Coal price)	2.61	(0.22)***	2.61	(0.22)***	2.64	(0.23)***
PTC*after	0.11	(0.08)	0.17	(0.08)*	0.16	(0.08)*
PTC	0.14	(0.06)**				
after	-0.26	(0.13)*	-0.29	(0.12)**	-0.39	(0.12)***
ln(Oil price)					0.21	(0.12)*
constant	-7.54	(0.66)***	-7.46	(0.67)***	-8.18	(0.80)***
Exponential time trends	Yes		Yes		Yes	
County fixed effects	No		Yes		Yes	
Counties			16		16	
Observations	239		239		239	
Instruments Joint F	68.26		71.86		72.40	

This table reports OLS regressions of the log of wood chip prices on the instruments and exogenous covariates. Standard errors are clustered at the county level. The data are a 15-year panel created from Maine Forest Service annual reports. ‘Coal Price’ is the annual average price of coal. ‘PTC\*after’ is an indicator for counties near biomass power plants in years the Producer Tax Credit was available. ‘PTC’ is an indicator for counties near biomass power plants. ‘After’ is a dummy for years after introduction of the PTC. All regressions include separate time trends for counties with and without the PTC. Legend for significance stars: \* 10%; \*\* 5%; \*\*\* 1%.

Table 3.5: Main results: The effect of wood chip price on wood chip quantity supplied

Dependent variable:  $\ln(\text{wood chip quantity per square mile})$

	OLS		IV, both		IV, coal only	
$\ln(\text{Price})$	0.23	(0.11)*	0.43	(0.23)*	0.64	(0.24)***
$\ln(\text{Oil Price})$	0.20	(0.12)	-0.15	(0.15)	0.03	(0.14)
after			0.40	(0.19)**		
Exponential time trends	Yes		Yes		Yes	
County fixed effects	Yes		Yes		Yes	
Counties	16		16		16	
Observations	239		239		239	
Overidentification (p-value)			0.86			

This table reports regressions of the log of wood chip quantity on log prices. Standard errors are clustered at the county level. “IV, both” instruments with coal price and the PTC subsidy. This regression includes separate exponential time trends for treatment and control counties with respect to the PTC. “IV, coal only” instruments with coal price. The overidentification p-value is the p-value for Hansen’s J statistic. The data are a 15-year panel created from Maine Forest Service annual reports. ‘Coal Price’ is the annual average price of coal. ‘PTC\*after’ is an indicator for county/year observations where the Producer Tax Credit was available. ‘After’ is a dummy for years after introduction of the PTC. Legend for significance stars: \* 10%; \*\* 5%; \*\*\* 1%.

## References

- Adams, William, Liran Einav, and Jonathan Levin (2009). “Liquidity Constraints and Imperfect Information in Subprime Lending”. *American Economic Review* 99 (1), pp. 49–84.
- Alberini, Anna and David Austin (2002). “Accidents Waiting to Happen: Liability Policy and Toxic Pollution Releases”. *The Review of Economics and Statistics* 84 (4), pp. 729–741.
- Allcott, H. and M. Greenstone (2012). “Is There an Energy Efficiency Gap?” *Journal of Economic Perspectives* 26 (1), pp. 3–28.
- Allcott, H., S. Mullainathan, and D. Taubinsky (2014). “Energy Policy with Externalities and Internalities”. *Journal of Public Economics* 112, pp. 72–88.
- Allcott, Hunt (2015). “Site Selection Bias in Program Evaluation”. *The Quarterly Journal of Economics*.
- Allcott, Hunt and Daniel Keniston (2014). *Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America*. Working Paper 20508. National Bureau of Economic Research.
- Alquist, Ron and Lutz Kilian (2010). “What do we learn from the price of crude oil futures?” *Journal of Applied Econometrics* 25 (4), pp. 539–573.
- Anderson, Owen L (1989). “The Anatomy of an Oil and Gas Drilling Contract”. *Tulsa Law Journal* 25, p. 359.
- Arimura, T. et al. (2012). “Cost-Effectiveness of Electricity Energy Efficiency Programs”. *Energy Journal* 33 (2), pp. 63–99.
- Auffhammer, M., C. Blumstein, and M. Fowlie (2008). “Demand-Side Management and Energy Efficiency Revisited”. *Energy Journal* 29 (3), pp. 91–104.
- Bain, Joe S. (1956). *Barriers to new competition: their character and consequences in manufacturing industries*. Harvard University Press Cambridge, MA.
- Baird, S. et al. (2012). “Designing Experiments to Measure Spillover and Threshold Effects”. *UC San Diego Working Paper*.

- Banerjee, Abhijit V and Esther Duflo (2005). “Growth theory through the lens of development economics”. *Handbook of Economic Growth* 1, pp. 473–552.
- Battistin, E. et al. (2009). “The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach”. *American Economic Review* 99 (5), pp. 2209–26.
- Baum-Snow, N. and J. Marion (2009). “The Effects of Low Income Housing Tax Credit Developments on Neighborhoods”. *Journal of Public Economics* 93 (56), pp. 654–666.
- Benjamin, Jeffrey G (2010). *Considerations and recommendations for retaining woody biomass on timber harvest sites in Maine*. Maine Agricultural and Forest Experiment Station.
- Benjamin, Jeffrey G., Robert J. Lilieholm, and Charles E. Coup (2010). “Forest Biomass Harvesting in the Northeast: A Special-Needs Operation?” *Northern Journal of Applied Forestry* 27 (2), pp. 45–49.
- Benneer, L., J. Lee, and L. Taylor (2013). “Municipal Rebate Programs for Environmental Retrofits: An Evaluation of Additionality and Cost-Effectiveness”. *Journal of Policy Analysis and Management* 32 (2), pp. 350–372.
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012). “The Organization of Firms Across Countries”. *The Quarterly Journal of Economics* 127 (4), pp. 1663–1705.
- Bovenberg, A. and L. Goulder (1996). “Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses”. *American Economic Review* 86 (4), pp. 985–1000.
- (2002). “Environmental Taxation and Regulation”. *Handbook of Public Economics* 3, pp. 1471–1545.
- Boyd, James (2002). “Financial responsibility for environmental obligations: An analysis of environmental bonding and assurance rules”. *Law and Economics of Environmental Policy, Research in Law and Economics Series, Elsevier*.
- Boyd, James and Daniel E Ingberman (2003). “Fly by night or face the music? Premature dissolution and the desirability of extended liability”. *American Law and Economics Review* 5 (1), pp. 189–232.
- Briedis, Julia I et al. (2011). “Biomass retention following whole-tree, energy wood harvests in central Maine: adherence to five state guidelines”. *Biomass and Bioenergy* 35 (8), pp. 3552–3560.
- Brooks, Richard RW (2002). “Liability and Organizational Choice”. *Journal of Law and Economics* 45 (1), pp. 91–125.
- Browning, Edgar K. (1987). “On the Marginal Welfare Cost of Taxation”. *American Economic Review* 77 (1), pp. 11–23.

- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2011). “Robust inference with multiway clustering”. *Journal of Business & Economic Statistics* 29 (2).
- Chandra, A., S. Gulati, and M. Kandlikar (2010). “Green Drivers or Free Riders? An Analysis of Tax Rebates for Hybrid Vehicles”. *Journal of Environmental Economics and Management* 60 (2), pp. 78–93.
- Che, Yeon-Koo and Kathryn E. Spier (2008). “Strategic judgment proofing”. *The RAND Journal of Economics* 39 (4), pp. 926–948.
- Commission for Environmental Cooperation (2011). “North American Power Plant Emissions”. *www.cec.org*.
- Darrah, Thomas H. et al. (2014). “Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales”. *Proceedings of the National Academy of Sciences* 111 (39), pp. 14076–14081.
- Davis, L., A. Fuchs, and P. Gertler (2014). “Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico”. *American Economic Journal: Economic Policy*.
- Davis, Lucas W. (2015). “Policy Monitor Bonding Requirements for U.S. Natural Gas Producers”. *Review of Environmental Economics and Policy* 9 (1), pp. 128–144.
- De Janvry, A. and E. Sadoulet (2006). “Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality”. *The World Bank Economic Review* 20 (1), pp. 1–29.
- Department of Energy (2011). *U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*. Oak Ridge National Laboratory.
- Einav, Liran, Mark Jenkins, and Jonathan Levin (2013). “The impact of credit scoring on consumer lending”. *The RAND Journal of Economics* 44 (2), pp. 249–274.
- European Commission Directorate General For Energy (2011). *Results of the Public Consultation on Additional Sustainability Measures at EU Level for Solid and Gaseous Biomass Used in Electricity, Heating, and Cooling*. European Commission.
- Evans, A.M. and R.T. Perschel (2009). *An assessment of biomass harvesting guidelines*. The Forest Guild.
- Feldstein, M. and C. Clotfelter (1976). “Tax Incentives and Charitable Contributions in the United States: A Microeconomic Analysis”. *Journal of Public Economics* 5 (12), pp. 1–26.
- Filmer, D. and N. Schady (2011). “Does More Cash in Conditional Cash Transfer Programs Always Lead to Larger Impacts on School Attendance?” *Journal of Development Economics* 96 (1), pp. 150–157.

- Gallagher, K. and E. Muehlegger (2011). "Giving Green to Get Green? Incentives and Consumer Adoption of Hybrid Vehicle Technology". *Journal of Environmental Economics and Management* 61 (1), pp. 1–15.
- Gan, Jianbang and CT Smith (2006). "A comparative analysis of woody biomass and coal for electricity generation under various CO<sub>2</sub> emission reductions and taxes". *Biomass and Bioenergy* 30 (4), pp. 296–303.
- Ganuza, Juan José and Fernando Gomez (2011). "Soft Negligence Standards and the Strategic Choice of Firm Size". *The Journal of Legal Studies* 40 (2), pp. 439–466.
- Gerard, David (2000). "The law and economics of reclamation bonds". *Resources Policy* 26 (4), pp. 189–197.
- Gerard, David and Elizabeth J Wilson (2009). "Environmental bonds and the challenge of long-term carbon sequestration". *Journal of Environmental Management* 90 (2), pp. 1097–1105.
- Goulder, Lawrence H., Ian W. H. Parry, and Dallas Burtraw (1997). "Revenue-Raising versus Other Approaches to Environmental Protection: The Critical Significance of Preexisting Tax Distortions". *The RAND Journal of Economics* 28 (4), pp. 708–731.
- Greenstone, M., E. Kopits, and A. Wolverton (2013). "Developing a Social Cost of U.S. Regulatory Analysis: A Methodology and Interpretation". *Review of Environmental Economics and Policy*.
- Greenstone, Michael and Justin Gallagher (2008). "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program". *The Quarterly Journal of Economics* 123 (3).
- Grosche, P. and C. Vance (2009). "Willingness to Pay for Energy Conservation and Free Ridership on Subsidization: Evidence from Germany". *The Energy Journal* 30 (2), pp. 135–154.
- Hahn, J., P. Todd, and W. Van der Klaauw (2001). "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design". *Econometrica* 69 (1), pp. 201–209.
- Harkins, Robert M (1994). "Lender Liability under Superfund: New Rules-Old Rules-No Rules". *Environs: Environmental Law & Policy Journal* 17, p. 18.
- Hartman, R. (1988). "Self-Selection Bias in the Evolution of Voluntary Energy Conservation Programs". *Review of Economics and Statistics*, pp. 448–458.
- Hsieh, Chang-Tai and Benjamin A. Olken (2014). "The Missing 'Missing Middle'". *Journal of Economic Perspectives* 28 (3), pp. 89–108.

- Innovative Natural Resource Solutions (2005). *Maine Future Forest Economy Project: Current Conditions and Factors Influencing the Future of Maine's Forest Products Industry*. Maine Forest Service.
- Ito, Koichiro (Forthcoming). "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program". *American Economic Journal: Economic Policy*.
- Jaffee, Dwight M. and Thomas Russell (1976). "Imperfect Information, Uncertainty, and Credit Rationing". *The Quarterly Journal of Economics* 90 (4), pp. 651–666.
- Janowiak, Maria K and Christopher R Webster (2010). "Promoting ecological sustainability in woody biomass harvesting". *Journal of forestry* 108 (1), pp. 16–23.
- Jones, Stephanie K. (2003). "Working With Blowouts, Knock-for-Knocks and Other Oil Patch Oddities". *Insurance Journal*.
- Joskow, P. and D. Marron (1992). "What Does a Negawatt Really Cost? Evidence from Utility Conservation Programs". *Energy Journal*, pp. 41–74.
- Kaiser, Mark J. and Brian Snyder (2009). "Raising supplemental bonding ups small company liabilities". *Oil and Gas Journal*.
- Kell, Scott (2011). *State Oil and Gas Agency Groundwater Investigations And Their Role in Advancing Regulatory Reforms-A Two-State Review: Ohio and Texas*.
- Kellogg, Ryan (2011). "Learning by drilling: Interfirm learning and relationship persistence in the Texas oilpatch". *The Quarterly Journal of Economics*, p. 39.
- Kornhauser, Lewis A (1982). "An economic analysis of the choice between enterprise and personal liability for accidents". *California Law Review*, pp. 1345–1392.
- Lee, D. and T. Lemieux (2010). "Regression Discontinuity Designs in Economics". *Journal of Economic Literature* 48, pp. 281–355.
- Li, Shanjun, Joshua Linn, and Erich Muehlegger (forthcoming). "Gasoline Taxes and Consumer Behavior". *American Economic Journal: Economic Policy*.
- Liao, Xianchun and Yaoqi Zhang (2008). "An econometric analysis of softwood production in the US South: a comparison of industrial and nonindustrial forest ownerships". *Forest Products Journal* 58 (11), pp. 69–74.
- Lilieholm R.J., P.R. Lammert G.R. Lorg and S.N. Trospen (2010). *Directory of Maine's Primary Wood Processors*. Maine Forest Service Department of Conservation.
- LoPucki, Lynn M (1996). "The death of liability". *Yale Law Journal*, pp. 1–92.
- Loughran, D. and J. Kulick (2004). "Demand-Side Management and Energy Efficiency in the United States". *Energy Journal*, pp. 19–43.



- Lucas, Robert E (1978). “On the size distribution of business firms”. *The Bell Journal of Economics*, pp. 508–523.
- Maine Forest Service (2008). *Assessment of Sustainable Biomass Availability*.
- Manomet Center for Conservation Sciences (2010). *Biomass Sustainability and Carbon Policy Study: Report to the Commonwealth of Massachusetts Department of Energy Resources*. Natural Capital Initiative Report NCL-2010-03.
- McCrary, J. (2008). “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test”. *Journal of Econometrics* 142 (2), pp. 698–714.
- McKinsey and Company (2009). *Unlocking Energy Efficiency in the U.S. Economy*.
- Mian, A. and A. Sufi (2012). “The Effects of Fiscal Stimulus: Evidence from the 2009 Cash for Clunkers Program”. *Quarterly Journal of Economics* 127 (3), pp. 1107–1142.
- Miguel, Edward and Michael Kremer (2004). “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities”. *Econometrica* 72 (1), pp. 159–217.
- Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins (2014). *The Housing Market Impacts of Shale Gas Development*. Working Paper 19796. National Bureau of Economic Research.
- Muller, N., R. Mendelsohn, and W. Nordhaus (2011). “Environmental Accounting for Pollution in the United States Economy”. *American Economic Review* 101 (5), pp. 1649–1675.
- Murray, Milissa A and Sandra Franco (2011). “Treatment of Environmental Liabilities in Bankruptcy”. In: *Environmental Aspects of Real Estate and Commercial Transactions*. Ed. by James B Witkin. American Bar Association.
- National Research Council (2013). *Alternatives for Managing the Nation’s Complex Contaminated Groundwater Sites*. Committee on Future Options for Management in the Nation’s Subsurface Remediation Effort.
- National Academy of Sciences, National Academy of Engineering, and National Research Council (2010). *Real Prospects for Energy Efficiency in the United States*. The National Academies Press.
- Newman, David H and David N Wear (1993). “Production economics of private forestry: a comparison of industrial and nonindustrial forest owners”. *American Journal of Agricultural Economics* 75 (3), pp. 674–684.
- North Carolina Clean Energy Technology Center (2012). *Database of State Incentives for Renewables & Efficiency*.
- Northeast Energy Solutions (2008). *Reliability Assessment Power Market Cost/Benefit Assessment of Maine Power Connection*.

- Parry, I. (1998). “A Second-Best Analysis of Environmental Subsidies”. *International Tax and Public Finance* 5 (2), pp. 153–170.
- Perlack R.D., L.L. Wright A.F. Turhollow R.L. Graham B.J. Stokes and D.C. Erbach (2005). *Biomass as Feedstock for a Bioenergy and Bioproducts Industry: The Feasibility of a Billion-Ton Annual Supply*. ORNL/TM-2005/66. Oak Ridge, TN: Oak Ridge National Laboratory.
- Pinkerton, John E (2007). “Sulfur dioxide and nitrogen oxides emissions from US pulp and paper mills, 1980–2005”. *Journal of the Air & Waste Management Association* 57 (8), pp. 901–906.
- Pitchford, Rohan (1995). “How liable should a lender be? The case of judgment-proof firms and environmental risk”. *The American Economic Review*, pp. 1171–1186.
- Polborn, Mattias K (1998). “Mandatory insurance and the judgment-proof problem”. *International Review of Law and Economics* 18 (2), pp. 141–146.
- Price Waterhouse Coopers (2011). *Global Forest, Paper, and Packaging Industry Survey. Survey of 2010 Results*.
- Rauch, James E. (1991). “Modelling the informal sector formally”. *Journal of Development Economics* 35 (1), pp. 33–47.
- Revelt, D. and K. Train (1998). “Mixed Logit with Repeated Choices: Households’ Choices of Appliance Efficiency Level”. *Review of Economics and Statistics* 80 (4), pp. 647–657.
- Ringleb, Al H and Steven N Wiggins (1990). “Liability and large-scale, long-term hazards”. *Journal of Political Economy*, pp. 574–595.
- Rothschild, Michael and Joseph Stiglitz (1976). “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information”. *The Quarterly Journal of Economics* 90 (4), pp. 629–649.
- Sallee, J. (2011). “The Surprising Incidence of Tax Credits for the Toyota Prius”. *American Economic Journal: Economic Policy* 3 (2), pp. 189–219.
- Sallee, J. and J. Slemrod (2012). “Car Notches: Strategic Automaker Responses to Fuel Economy Policy”. *Journal of Public Economics* 96, pp. 981–999.
- Schneider, L. (2007). “Is the CDM Fulfilling its Environmental and Sustainable Development Objectives? An Evaluation of the CDM and Options for Improvement”.
- Shavell, Steven (1984). “A model of the optimal use of liability and safety regulation”. *The Rand Journal of Economics* 15 (2), pp. 271–280.
- (1986). “The Judgment Proof Problem”. *International Review of Law and Economics* 6 (1), pp. 45–58.

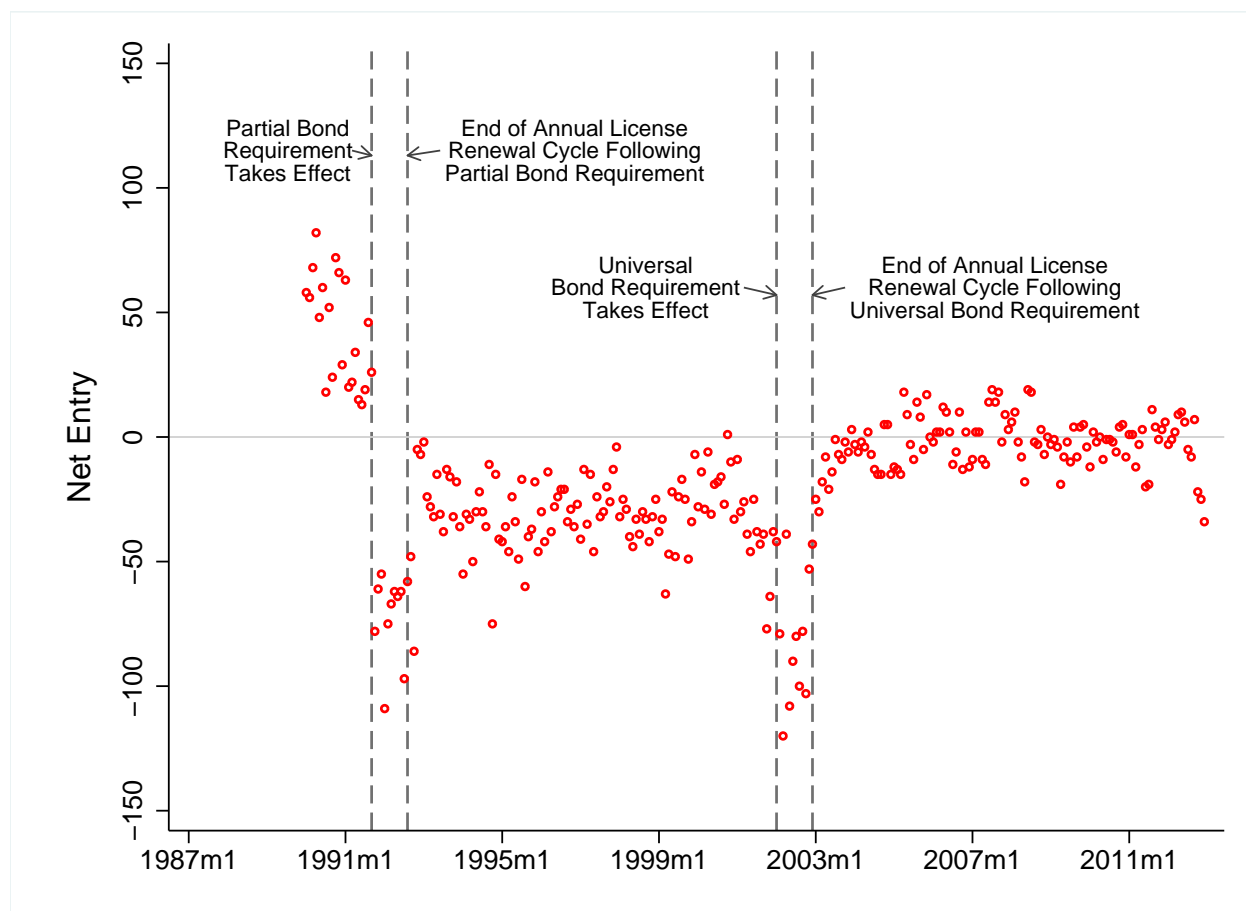
- Shavell, Steven (2002). *Minimum asset requirements*. Working Paper 9335. National Bureau of Economic Research.
- (2005). “Minimum Asset Requirements and Compulsory Liability Insurance as Solutions to the Judgment-Proof Problem”. *RAND Journal of Economics*, pp. 63–77.
- (2007). “Chapter 2 Liability for Accidents”. In: ed. by A.M. Polinsky and S. Shavell. Vol. 1. *Handbook of Law and Economics*. Elsevier, pp. 139–182.
- Sinai, T. and J. Waldfogel (2005). “Do Low-Income Housing Subsidies Increase the Occupied Housing Stock?” *Journal of Public Economics* 89 (1112), pp. 2137–2164.
- Snchez-Azofeifa, G.A. et al. (2007). “Costa Rica’s Payment for Environmental Services Program: Intention, Implementation, and Impact”. *Conservation Biology* 21 (5), pp. 1165–1173.
- State of Maine (2008). *Governor’s Wood-to-Energy Task Force Report*.
- STRONGER (2003). *State Review of Oil and Natural Gas Environmental Regulations: Texas State Review*. Environmental Protection Agency, Interstate Oil, and Gas Compact Commission.
- Texas House of Representatives (2002). *House Committee on Environmental Regulation Interim Report 2002*.
- U.S. IAWG (2013). *Technical Support Document: Technical Update on the Social Cost of Carbon for Regulatory Impact Analysis*. U.S. Interagency Working Group on the Social Cost of Carbon for Regulatory Impact Analysis.
- Van ’t Veld, Klaas (2006). “Hazardous-industry restructuring to avoid liability for accidents”. *International Review of Law and Economics* 26, pp. 297–322.
- Viner, Jacob (1932). *Cost curves and supply curves*. Springer.
- White, Michelle J. (2007). “Chapter 14: Bankruptcy Law”. In: ed. by A.M. Polinsky and S. Shavell. Vol. 2. *Handbook of Law and Economics*. Elsevier, pp. 1013–1072.
- Wiltsee, G. (2000). *Lessons Learned from Existing Biomass Power Plants*. Report NREL/SR-570-26946. National Renewable Energy Laboratory.
- Wiser R., Namovicz C. Gielecki M. and R. Smith (2007). *Renewables Portfolio Standards: A Factual Introduction to Experience from the United States*. LBNL-62569. Lawrence Berkeley National Laboratory Environmental Energy Technologies Division.
- Wiser, R. (2010). *State of the States: Update on RPS Policies and Progress. Presentation at Renewable Energy Markets 2010. October 20, 2010. Portland, OR*. Lawrence Berkeley National Laboratory Environmental Energy Technologies Division.

- Wolfram., C, O. Shelef, and P. Gertler (2012). “How Will Energy Demand Develop in the Developing World?” *Journal of Economic Perspectives* 26, pp. 119–138.
- Working, Elmer J (1927). “What Do Statistical “Demand Curves” Show?” *The Quarterly Journal of Economics*, pp. 212–235.
- Yin, Haitao, Alex Pfaff, and Howard Kunreuther (2011). “Can Environmental Insurance Succeed Where Other Strategies Fail? The Case of Underground Storage Tanks”. *Risk Analysis* 31 (1), pp. 12–24.

## Appendix A

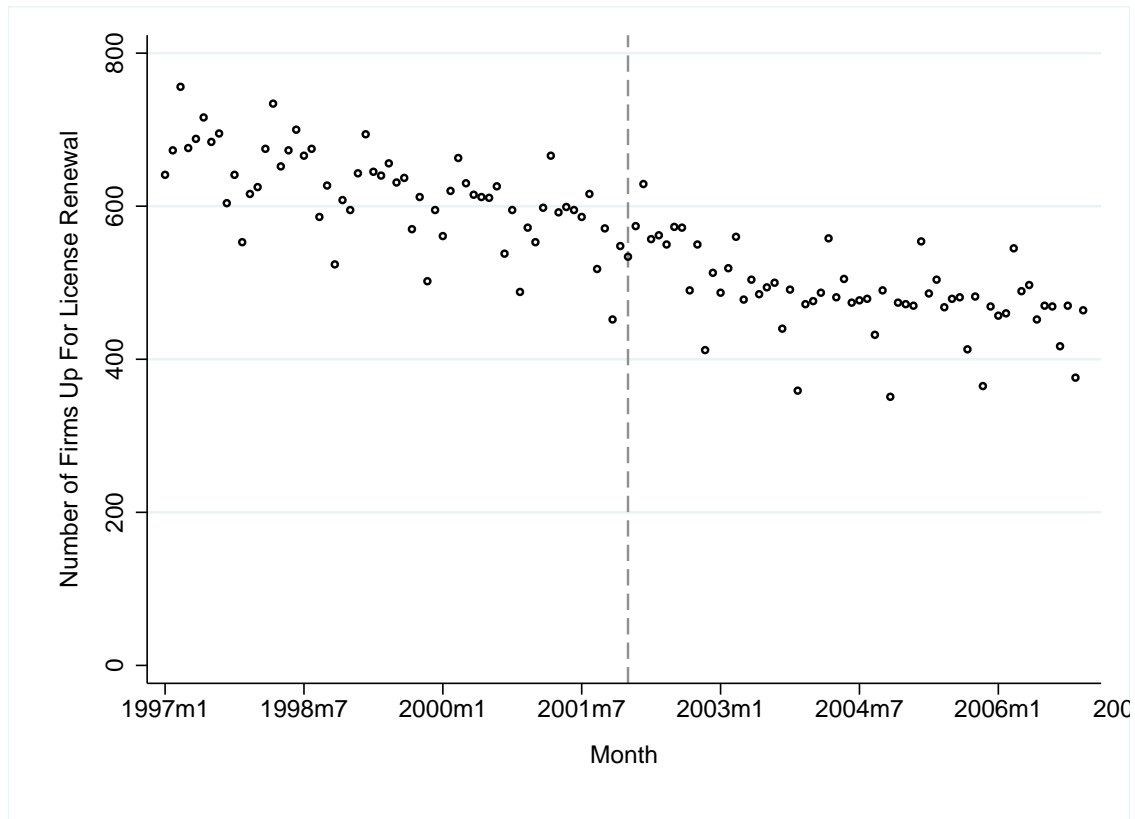
# Bankruptcy and Environmental Risk Appendix

Figure A.1: Net Entry (Entry - Exit) by Month



Notes: This figure shows the net change in the number of firms in the industry each month. Entry date is defined as the date of the firm's first annual operating license filing. Exit date is defined as 365 days after the firm's final annual operating license renewal. The sample includes all firms with oil or gas leases from 1990 to 2010. The initial rollout of the partial bond requirement occurred during September, 1991 to August, 1992. The initial rollout of the universal bond requirement occurred during January–December, 2002.

Figure A.2: Number of Firms Up For License Renewal Each Month



Notes: This figure shows the number of firms required to renew their annual operating license (or exit the industry) in each month. These months are assigned by the Texas Railroad Commission and cannot be manipulated by firms. The sample includes firms with oil or gas production during 1997–2006. The vertical dashed line indicates the implementation of the increased bond requirement in January 2002.

Figure A.3: Number of Orphan Wells By Operator's Quarter of Exit

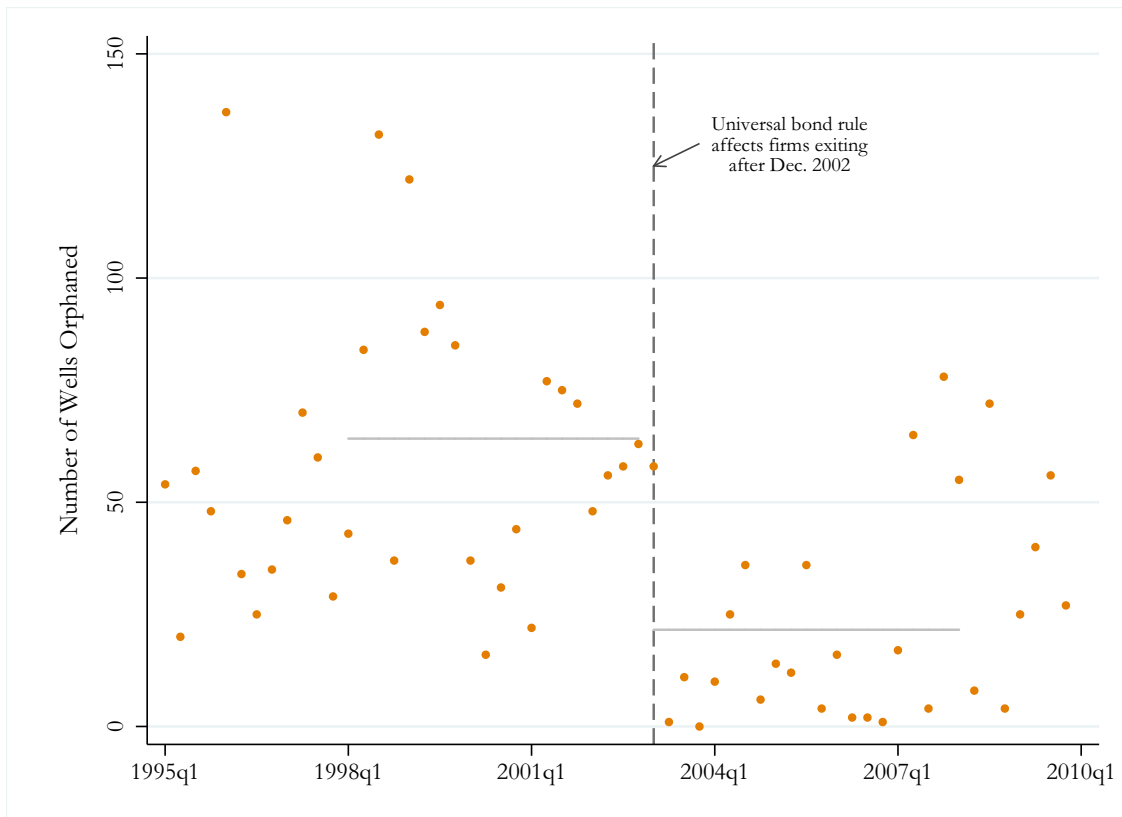
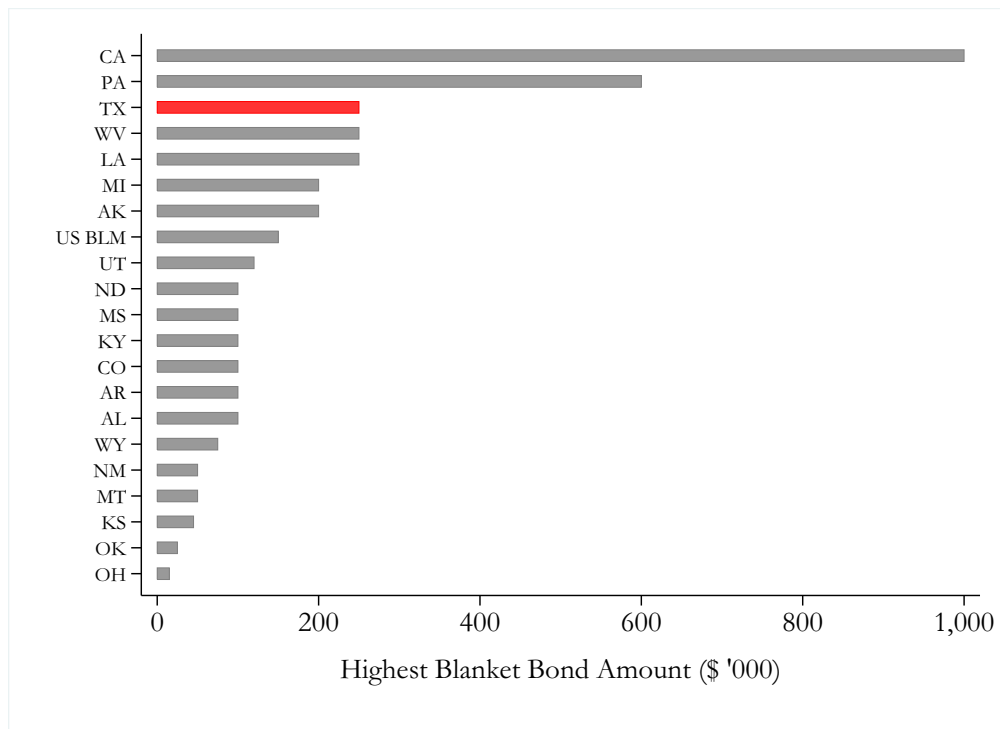




Figure A.4: Bond Requirements By State



Notes: This figure shows the maximum required bond per firm in the 20 U.S. states with the most drilling activity in 2012, and the federal Bureau of Land Management. Information on bond requirements is from Penn Environment Center 2013. "Who Pays the Cost of Fracking?" <http://pennenvironmentcenter.org>.

Table A.1: Comparing Firms by License Renewal Month

Month	(1) Number of Firms	(2) Mean Annual Production (\$)	(3) Std. Dev. Annual Production (\$)
January	769	1,450,156	5,388,415
February	791	1,409,704	5,358,007
March	874	1,451,261	5,476,291
April	784	1,156,730	4,729,774
May	760	1,157,181	4,271,193
June	807	1,429,897	5,545,757
July	763	1,037,615	4,085,928
August	773	1,165,920	4,312,536
September	673	1,414,795	5,316,751
October	734	1,001,681	3,812,794
November	609	1,334,588	4,806,508
December	684	1,487,344	5,198,563
F statistic		0.97	
p-value		0.47	

Notes: This table reports summary statistics for firms according to the month which contains their annual license renewal date. This table covers firms from their 1996 to 2000 license renewals. Annual production is twelve times the average value of monthly oil and natural gas production, calculated using oil and gas prices in each month. To reduce the noise caused by a few very large firms, I drop firms larger than the 99th percentile of annual average production value (for this table only). The F statistic and p-value are for a test of the null hypothesis that the mean of average annual production is the same in every group.

Table A.2: Controlling Separately for Crude Oil and Natural Gas Prices

	(1)	(2)	(3)	(4)	(5)
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Begin Rollout	0.110*** (0.017)	0.122*** (0.028)	0.046** (0.019)	0.047*** (0.014)	-0.018 (0.014)
End Rollout	-0.153*** (0.032)	-0.114*** (0.022)	-0.091*** (0.019)	-0.050** (0.019)	-0.011 (0.015)
Oil Price (\$100/bbl)	-0.080 (0.057)	-0.088** (0.033)	-0.050 (0.036)	-0.037 (0.029)	0.047 (0.031)
Natural Gas Price (\$100/mcf)	-0.508 (0.354)	0.072 (0.272)	-0.436** (0.179)	0.095 (0.166)	-0.331** (0.139)
Constant	0.177*** (0.015)	0.093*** (0.022)	0.085*** (0.013)	0.049*** (0.009)	0.102*** (0.010)
N	11,903	13,267	14,667	16,197	15,792
Firms	2,129	2,185	2,202	2,226	2,236

Notes: This table is identical to table 1.3, except that crude oil and natural gas prices, which are highly correlated during this period, are both included in the regressions.

Table A.3: Texas Railroad Commission Regulations Implemented in 2001 and 2002

Action	Proposed	Implemented
Allows electronic filing of drilling permits	March 2001 (TXR 26 2257)	June 2001 (TXR 26 4088)
Clarifies wording of hazardous waste rules <sup>1</sup>	May 2001 (TXR 26 3431)	September 2001 (TXR 26 6870)
Extends existing tax credit for high-cost gas	June 2001 (TXR 26 4015)	August 2001 (TXR 26 6009)
Extends existing tax credit for marginal wells	July 2001 (TXR 26 3431)	September 2001 (TXR 26 6869)
<b>Implements Senate Bill 310 (bonding)<sup>2</sup></b>	<b>August 2001</b> (TXR 26 5919)	<b>January 2002</b> (TXR 27 139)
Clarifies rules for assigning acreage to pooled units	October 2001 (TXR 26 7721)	January 2002 (TXR 27 150)
Clarifies rules for requesting end to unitization	November 2001 (TXR 26 9480)	February 2002 (TXR 27 906)
Clarifies rules for transporting oil and gas	January 2002 (TXR 27 547)	May 2002 (TXR 27 3756)
Clarifies rules for “swabbing” existing wells <sup>3</sup>	April 2002 (TXR 27 2666)	September 2002 (TXR 27 9149)

Notes: This table lists all rules changes for oil and gas producers implemented by the Texas Railroad Commission during 2001 and 2002. It is based on all rule introductions or amendments listed in the RRC Oil and Gas Division rules (Texas Administrative Code, Title 16, Part 1, Chapter 3). “TXR” refers to volume and page number in the Texas Register. The date proposed is the date that the rule was published as a “Proposed Rule” to allow for public comment. The date implemented is the date that the regulation was published as an “Adopted Rule”.

<sup>1</sup>This was a technical change in wording to match federal law, changing the word “facility” to “site.” The proposed rule states, “The language change is consistent with the way the commission has applied the rule in that the commission’s intent and policy, since the initial adoption of 3.98 in 1996, has been to apply the provisions of subsection (e) to oil and gas waste generators. Therefore, no one will be affected that was not affected under the previous rule.”

<sup>2</sup>SB 310 passed the Texas legislature in June 2001; the RRC rule implementing SB 310 was first published as a proposed rule in August, 2001. This version was withdrawn and a second proposed rule was published in November, 2001 (TXR 26 8937).

<sup>3</sup>Swabbing is a technique that involves pulling fluid through the well bore using a wire and cup assembly. This rule clarifies that swabbing is prohibited as an ongoing production method to extend the life of very old wells.

## Appendix B

### Additionality in Energy Efficiency Appendix

## Calculating Historical Consumption

This appendix describes the program rules that were used in calculating average historical electricity consumption for each household. As described in the text, participating retailers determined subsidy amounts for households using a specially-designed website. In determining the subsidy amount, the website calculated the average historical electricity consumption for each household as of the date they were applying. These calculations were not archived, but we have used our database of electric billing records to recreate these calculations as accurately as possible.

Before making these calculations we first merged the program data with the database of electric billing records. For 86% of program participants, there was a household with the identical account number in the electric billing records. We exclude the remaining 14% for whom there is no perfect match, leaving us with 998,930 total program participants. We also drop a small number of records (less than one-tenth of one percent) which were improperly formatted.

Using billing cycle codes, we determined as accurately as possible the exact days corresponding to all 300+ million billing cycles in our data. This was important because we needed to determine which bills were in the system as of each potential date of application. Nationwide 93% of residential electricity customers are billed in two-month cycles. Half of these households have their meters read during odd-numbered months (January, March, etc.) and half have their meters read during even-numbered months. An additional 5% of households have their meters read every month. The remaining 2% have billing periods of 3 months or longer. These irregular periods arise for a variety of reasons. For example, some households in extremely rural areas have their meters read less than six times per year.

It is important for our RD analysis that we recreate the eligibility calculations as closely as possible. The calculation is simplest and most transparent for households who were billed regularly throughout the sample period. Therefore, we drop 6,028,939 households (23%) where the billing cycle length is not constant or is greater than two months. For the refrigerator analysis, because exact billing dates are important (as described below), we drop an additional 2,417,350 households (9%) where the data do not provide the exact billing dates.

Program rules differed between air conditioners and refrigerators. For air conditioner replacements, average historical electricity consumption was calculated using all summer bills during the previous calendar year. “Summer” is defined as the six months of the year which have historically had the highest average temperature. This differs across locations. The four possibilities are February - July, March - August, April - September, and May - October.

For refrigerator replacements, the average was calculated using the most recent 12 months of bills that were available in the online eligibility system. In practice, it seems to have taken about 60 days on average between the day a meter was read and the day that information became available in the system. So, for example, a household whose meter was read in

odd-numbered months and who bought a refrigerator in May 2011 would have had their eligibility calculated using bills from April 2010 through March 2011.

For refrigerators the calculation also depended on the household's electricity rate structure. Residential customers in hot parts of the country have electric rates which vary seasonally, while customers in cool parts of the country have rates which are the same all year. For households whose electric rates do not vary seasonally, all of the bills during the most recent 12 months are included when calculating eligibility. For households whose electric rates vary seasonally, only the subset of non-summer bills during that period were used to calculate baseline consumption.

For participants, we calculate historical average electricity consumption on the day the appliance was purchased. For non-participants, we calculate historical average electricity consumption as of a date randomly chosen from the distribution of observed replacement dates in that non-participant's state.

For households participating in the program near the end of a billing cycle, a few days can change average historical consumption considerably depending on whether or not that last cycle was included in the average or not. In reconstructing average historical consumption for refrigerators we did the best we could, but we do not know exactly the date in which each billing cycle was included in the billing system, and thus cannot reconstruct this perfectly.

Appendix Table B.1 shows the number of total households and number of participants within our preferred bandwidths above or below each threshold. It also shows the number of participants within that range for whom we incorrectly predict historical usage. For air conditioners the predictions are extremely accurate, with correct predictions for 99% of participants at all thresholds. For refrigerators the predictions are less accurate but still very good. For refrigerators we correctly predict subsidies for at least 97% of participants at all thresholds. Moreover, the incorrect predictions are approximately symmetric, with near-equal numbers of participants receiving higher and lower subsidy amounts than we predict.

## Addressing Measurement Error

This section describes how we address measurement error in average historical electricity consumption. This variable is measured with error for a subset of observations, as described in the text. This is an issue only in the refrigerator program; in the air conditioner program there is essentially no measurement error. Eligibility for the air conditioner subsidies was calculated using summer consumption during the previous calendar year, regardless of when during the year the appliance was purchased. Eligibility for refrigerators, in contrast, was based on non-summer consumption during the most recent 12 months of billing history.

We observe the dates for each billing cycle, but not the dates when meter readings were entered into the central system. Consequently, it is impossible for us to be sure which bills were in the system at the moment a household participated in the program. In other words, we cannot be sure for each household which billing cycles were used when calculating eligibility for refrigerators. This means that for some observations we measure average historical consumption with error because we use a different set of billing cycles than those that were actually used. The direction and magnitude of this error will vary across households. Importantly, however, for most households we measure historical consumption with no error because we use the same billing cycles that were actually used to calculate eligibility.

For the set of observations where average historical electricity consumption is measured with error, the subsidy received will not change discontinuously at the observed threshold. Random noise in the running variable instead leads to a smooth “S”-shaped relationship between subsidy amount and historical average electricity consumption for these observations. Thus, for the full sample, the observed discontinuity in subsidy amounts at the threshold is reduced by the fraction of households for whom the running variable includes measurement error. At the three refrigerator thresholds, the observed change in the probability of receiving the higher subsidy ranges from 0.83 to 0.85.

(Battistin et al., 2009) shows that the fuzzy RD method consistently estimates the true treatment effect when the running variable is measured with error for only some observations, as long as the measurement error is independent of treatment and outcomes, conditional on the true value of the running variable. In our case, this is a reasonable assumption given that it seems unlikely that uncertainty about bill delivery dates would be correlated with a household’s electricity consumption. To formalize this, let  $X_{obs} = ZX^* + (1 - Z)X$ , where  $X_{obs}$  is observed historical average electricity consumption,  $X^*$  is true historical average electricity consumption, and  $X$  is equal to  $X^*$  plus random error. Let  $Z$  take the values 0 or 1 and  $Pr[Z = 1] = p$ . In other words,  $p$  is the probability that we observe the true value of historical average electricity consumption for any observation.

Let  $Y$  represent the binary outcome variable, program participation. As in Hahn, Todd, and Klaauw (2001), let  $Y^+ = \lim_{x \rightarrow 0^+} E[Y|X^* = x]$  and  $Y^- = \lim_{x \rightarrow 0^-} E[Y|X^* = x]$ . With no measurement error, the sharp RD estimator  $Y^+ - Y^-$  yields the true treatment effect  $\omega$ . However, with the type of measurement error described above, the sharp RD estimator is



biased downwards. Let  $\tilde{Y}^+ = \lim_{x \rightarrow 0^+} E[Y|X_{obs} = x]$  and  $\tilde{Y}^- = \lim_{x \rightarrow 0^-} E[Y|X_{obs} = x]$ . Battistin et al. (2009) shows that

$$\tilde{Y}^+ - \tilde{Y}^- = p \left[ \lim_{x \rightarrow 0^+} E[Y|X^* = x] - \lim_{x \rightarrow 0^-} E[Y|X^* = x] \right] = p\omega.$$

The sharp RD estimator underestimates the treatment effect  $\omega$  by a factor  $p$ .

Define  $\tilde{S}^+$  and  $\tilde{S}^-$  analogously for the treatment, receiving a higher subsidy offer. The true change in the probability of receiving the higher subsidy offer at the threshold is 1. But the estimator  $\tilde{S}^+ - \tilde{S}^-$  underestimates the change in probability of treatment by a factor  $p$ :

$$\tilde{S}^+ - \tilde{S}^- = p \left[ \lim_{x \rightarrow 0^+} E[S|X^*] - \lim_{x \rightarrow 0^-} E[S|X^*] \right] = p.$$

The fuzzy RD estimator  $\frac{\tilde{Y}^+ - \tilde{Y}^-}{\tilde{S}^+ - \tilde{S}^-}$  consistently estimates the true treatment effect by canceling out the downward bias  $p$ .

Accordingly, we estimate  $\tilde{S}^+ - \tilde{S}^-$  using the following first stage regression, as explained in the text:

$$1[\text{Larger Subsidy}]_i = \phi + g(X_i) + \gamma 1[\text{Below Threshold}]_i + \epsilon_i.$$

Because we do not observe what subsidy non-participants were (or would have been) offered, the first stage regression includes only the households within the chosen bandwidth who participated in the program. Because uncertainty about bill delivery dates is unlikely to be correlated with program participation, this first stage relationship is likely to be the same for participants as for the full population.

Finally, to calculate the fuzzy RD estimates of the increase in participation at the threshold, we divide the coefficient  $\rho$  from section 2.4 in the text by the estimated coefficient  $\gamma$  and we calculate standard errors using a block bootstrap with 5,000 repetitions. In addition to reporting the implied slope (change in share participating per dollar) and the implied price elasticities, we report the percentage change in participation.

For this last calculation, we divide this estimated increase in participation by  $\alpha$ , the share of households participating at the threshold without the larger subsidy. Ideally one would also adjust  $\alpha$  to correct for measurement error. We ignore this issue, however, and instead emphasize the slope estimates which do not require estimating  $\alpha$ . Moreover, our fuzzy RD estimates are very similar in magnitude to the sharp RD estimates, suggesting that failing to address measurement error in our estimates of  $\alpha$  only introduces a modest amount of bias into our estimates of the percentage change.

## Testing for Strategic Behavior

In this section we perform a final test aimed at a more subtle form of strategic behavior. In particular, we test whether households strategically *delayed* participation when they were close to a threshold. Suppose a household applies for the program and somehow learns that they just narrowly missed eligibility for one of the more generous subsidies. At least in theory, this household could wait for a billing cycle (or more), perhaps while simultaneously taking steps to reduce electricity consumption, and then reapply. This kind of strategic delay would lead us to find more participants just on the generous sides of these thresholds relative to a program for which eligibility was assigned only once.

We test for strategic delay for refrigerators but not air conditioners. Subsidy amounts for air conditioner replacement were based on summer consumption in the previous calendar year, so a household would have to wait much longer on average before changes in electricity consumption could affect their subsidy eligibility. To test for strategic delay in the refrigerator program, we examine the subsidies that households would have been eligible for one billing cycle before participating in the program. If there is no strategic delay, the probability of a household's average historical electricity consumption falling enough during this interval to increase their subsidy eligibility should be the same for participants and non-participants. In contrast, if participants are more likely than non-participants to be eligible for larger subsidies at purchase than sixty days earlier, this may indicate strategic delay.<sup>1</sup>

We implement this test using households within 15 kWh below the 175 kWh and 200 kWh thresholds (i.e., households who just barely qualified for the larger subsidy at each threshold). We calculate average historical electricity consumption for participants on the day that they participated, and for non-participants on a random date drawn from the empirical distribution of participation dates in each state. Then, for both groups, we also calculate average historical electricity consumption sixty days earlier.

Appendix Table B.2 reports the results. Overall, about one quarter of these households would have qualified for a less generous subsidy during the previous billing period. The similarity between participants and non-participants suggests that these changes were not driven by strategic delay. For households barely qualifying for the \$170 subsidy, 22.7% of participants and 22.6% of non-participants were eligible for smaller subsidies one billing period earlier. Using a two-sample t-test, we test the null hypothesis that these shares are equal for participants and non-participants. We find no statistically significant difference. For households barely qualifying for the \$110 subsidy, 23.2% of participants and 24.5% of non-participants would have been eligible for a smaller subsidy one billing period earlier.

---

<sup>1</sup>Another possible test would consider only participants. If there were no time-varying shocks to electricity consumption and no strategic delay, the share of smaller-subsidy recipients who were previously eligible for the larger subsidy should equal the share of larger-subsidy recipients who were previously eligible for the smaller subsidy. In practice, weather and other shocks complicate this comparison. For example, if the billing cycle before participation included many hot days, this would tend to increase consumption for all households and make decreases in subsidy eligibility more likely. Thus, we prefer the test that compares participants and non-participants.

The difference between these shares is small, but statistically significant at the 5% level. However, the effect is the opposite of what would be expected if participants were acting strategically: among this group, program participants were slightly *less* likely than non-participants to have experienced an increase in the subsidy for which they were eligible. In short, we find no evidence of strategic delay.

## Back-of-the-Envelope Welfare Calculation

### Electricity Savings Per Replacement

Electricity savings come from Johnson, Todd M., Claudio Alatorre, Zayra Romo, and Feng Lui. 2009. "Low-Carbon Development for Mexico," World Bank, Conference Edition (referred to as "Johnson 2009" from here on). We use the participation-weighted average of 481 kWh per year for refrigerators and 1,200 kWh per year for air conditioners. Following (Davis, Fuchs, and Gertler, 2014), we assume the program accelerates appliance replacement by five years. This yields an average lifetime electricity savings of 2,885 kWh per replacement.

### Greenhouse Gases

In Mexico 0.538 tons of carbon dioxide are emitted per megawatt hour of electricity generation, according to Johnson 2009, page 26. So 2,885 kWh of electricity consumption implies 1.55 tons of carbon dioxide.

(Greenstone, Kopits, and Wolverton, 2013) present a range of values for the social cost of carbon dioxide emissions according to different discount rates and for different time periods that is intended to capture changes in net agricultural productivity, human health, property damages from increased flood risk, and other factors. These estimates were then updated by (U.S. IAWG, 2013). With a 3% discount rate the central value for the social cost of carbon dioxide is \$34 per ton. Thus reducing carbon dioxide emissions by 1.55 tons generates benefits worth \$53.

### Local Air Pollution

We are not aware of any estimates of these benefits for Mexico, but (Muller, Mendelsohn, and Nordhaus, 2011) estimates the external damages from sulfur dioxide, nitrogen oxide, and particulates for different forms of U.S. power generation. Coal-fired power plants are the most damaging (2.8 cents per kilowatt hour), while oil (2.0 cents) and, in particular, natural gas (0.2 cents) are less damaging. Taking into account that Mexican plants emit higher levels of criteria pollutants than U.S. plants, and using the mix of electricity generation in Mexico, the estimates from Muller, Mendelsohn, and Nordhaus, 2011 imply that the benefits from reduced emissions of criteria pollutants are 1.9 cents per kilowatt-hour.

We calculated average emissions factors for Mexico and the United States using total electricity generation and emissions levels from Commission for Environmental Cooperation, 2011, Tables 2.2 and Table 2.4. Mexican plants emit per kilowatt hour 2.4 times as much sulfur dioxide, 1.7 times as much nitrogen oxide, and 2.2 times as much particulates ( $PM_{10}$ ). According to Muller, Mendelsohn, and Nordhaus, 2011 sulfur dioxide is the most damaging criteria pollutant so we scale damages by 2.4.

According to SENER, “Prospectiva del Sector Eléctrico, 2010-2025”, Figure 37, net generation in Mexico in 2010 was 52% natural gas, 17% oil, 12% coal, 11% hydroelectric, and 7% nuclear and non-hydro renewables. For criteria pollutants as well as for carbon dioxide emissions, what matters for benefits is *marginal* emissions, not average emissions. The two are likely similar in Mexico’s case, however, because generation is dominated by natural gas and neither coal nor hydro are likely to be marginal.

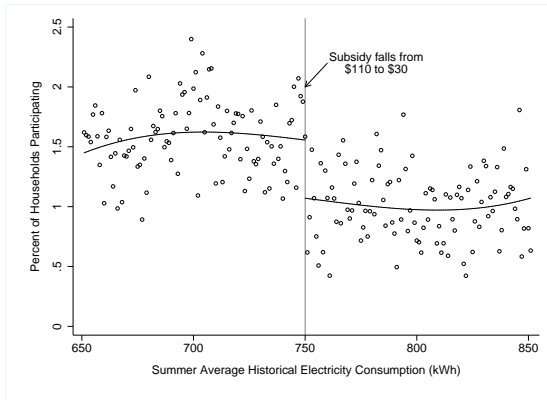
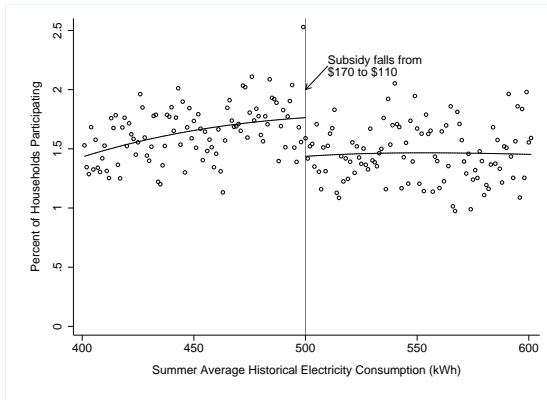
### Efficiency Costs of Subsidies Due to Pre-Existing Taxes

For this parameter we relied on studies of the United States because we are unaware of any studies of Mexico. The number 1.3 comes from a study of sulfur dioxide cap-and-trade regulation that considers both the revenue recycling and tax interaction effects (Goulder, Parry, and Burtraw, 1997). They find that the total cost of achieving the 1990 US Clean Air Act Amendment targets with a revenue-neutral cap-and-trade system was 1.3 times higher than in a first-best world with no pre-existing labor tax.

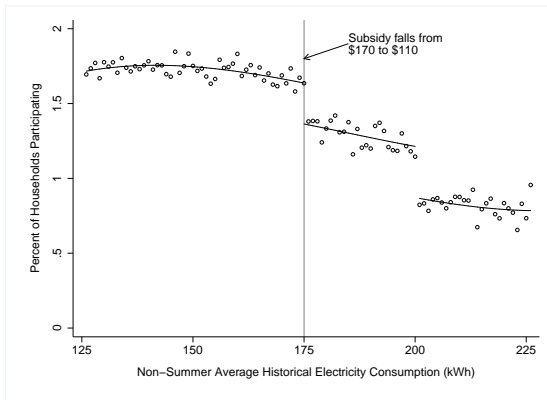
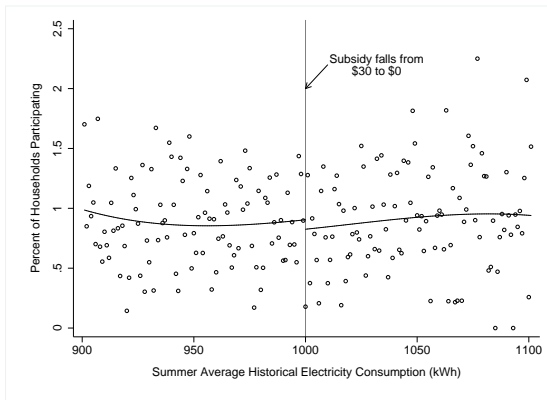
The number 1.3 is also consistent with Parry (1998), who finds that the net revenue-financing and tax interaction effects of environmental subsidies are greater than zero but below the costs implied by the revenue-financing effect alone. For subsidies that induce large percentage increases in consumption (like the one we are focused on), (Parry1998) finds that the difference between the revenue-financing effect and the net efficiency cost is small. Using estimates of the marginal cost of public funds (MCPF) from (Browning, 1987), the revenue-financing effect by itself would make the policy 1.32 to 1.47 times more expensive than in a world with no pre-existing distortions. Bovenberg and Goulder (2002) report a somewhat wider range.

Figure B.1: The Effect of Subsidy Size on Participation, RD Estimates

(a) Air Conditioners, 500 kWh Threshold (b) Air Conditioners, 750 kWh Threshold



(c) Air Conditioners, 1000 kWh Threshold (d) Refrigerators, 175 kWh Threshold



(e) Refrigerators, 200 kWh Threshold (f) Refrigerators, 250 kWh Threshold

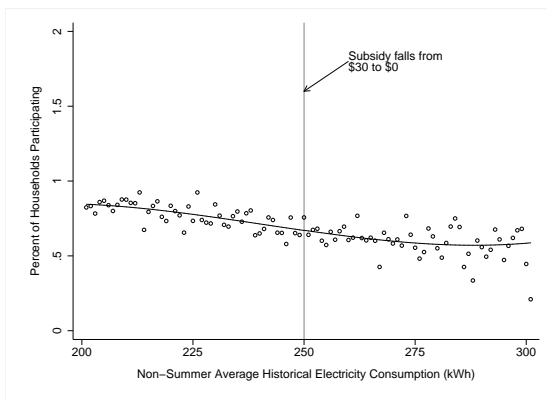
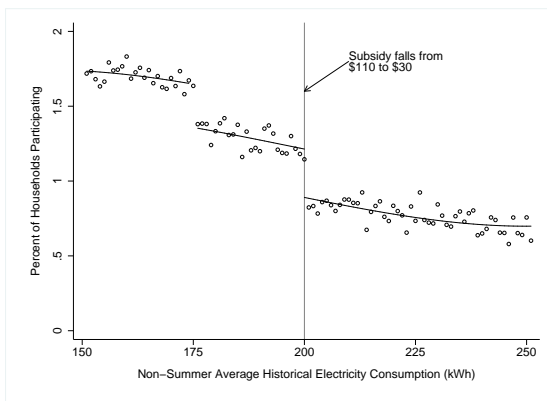


Table B.1: Number of Participants Close to Thresholds

Subsidy Increase	Number of Households Within Preferred Bandwidth	Number of Participants Within Preferred Bandwidth	Number of Participants Incorrectly Predicted to Receive Low Subsidy	Number of Participants Incorrectly Predicted to Receive High Subsidy	Percent Correctly Predicted Within Preferred Bandwidth
(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Air Conditioners					
\$110 to \$170	626,397	9,699	15	24	99.6
\$30 to \$110	241,218	3,213	7	10	99.5
\$0 to \$30	110,572	986	3	4	99.3
Panel B. Refrigerators					
\$110 to \$170	4,794,466	73,462	844	697	97.9
\$30 to \$110	3,218,317	41,953	487	307	98.1
\$0 to \$30	1,416,691	10,320	162	125	97.2

Notes: This table describes the number of households near each of the eligibility thresholds used in the RD analysis. In column (2), we report the number of households with average historical electricity consumption above or below each threshold by less than our preferred bandwidth of 100 kilowatt-hours for air conditioners and 50 kilowatt-hours for refrigerators. Column (3) gives the number of program participants in the same range. We then report in columns (4) and (5) the number of participants for whom our reconstruction of average historical electricity consumption incorrectly predicts the subsidy amount received. Column (6) reports the percentage of participants for whom we correctly predict the subsidy amount.

Table B.2: Testing for Strategic Delay

	Percent who would have been eligible for a smaller subsidy two months earlier	Percent who would have been eligible for the same or larger subsidy two months earlier	$\chi^2$ test, $H_0$ : identical distributions for participants and non-participants (p-value)
	(1)	(2)	(3)
Panel A. Households Barely Qualifying for \$170 Subsidy			
Participants	22.7%	77.3%	0.62
Non-participants	22.6%	77.5%	
Panel B. Households Barely Qualifying for \$110 Subsidy			
Participants	23.2%	76.8%	0.02
Non-participants	24.5%	75.5%	

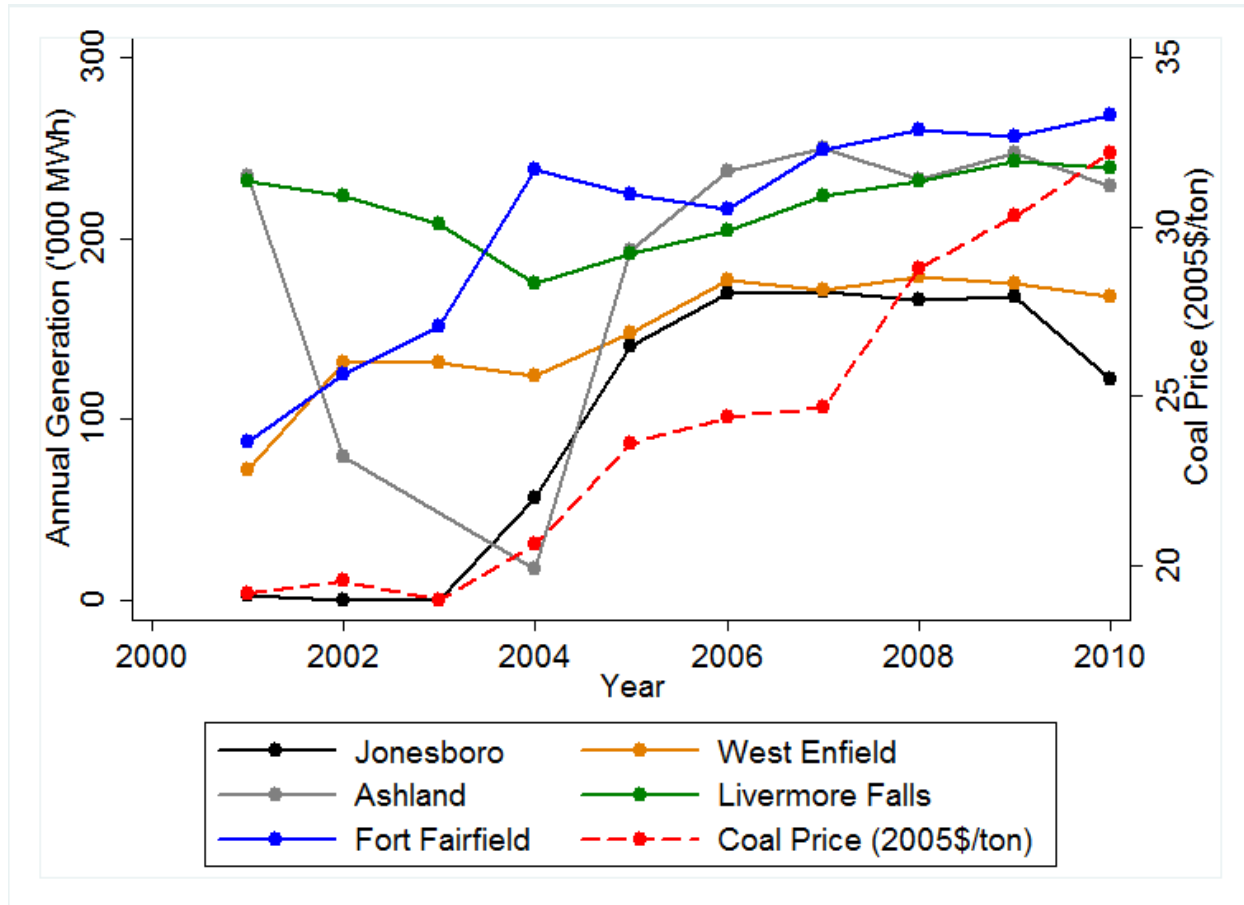
Notes: This table tests for strategic delay in the refrigerator replacement program. For households who just barely qualified for the \$170 and \$110 subsidies (i.e., within 15 kWh), we determine what subsidy each household would have qualified for one billing cycle earlier. See text for details.



## Appendix C

# Biomass Energy and Forests Appendix

Figure C.1: Annual generation from biomass plants increases as coal prices rise



The left axis of this figure shows total annual electricity generation at five biomass power plants in Maine. These data come from EIA Form 923. The right axis shows the annual average price of coal from the EIA Annual Energy Review. This figure suggests the possibility that rising coal prices may have contributed to additional operation of biomass power plants, perhaps by changing the order of marginal costs in the electricity industry supply curve.

Table C.1: Coal prices are uncorrelated with natural gas and housing starts

	Annual natural gas price	Annual housing starts
Annual coal price	0.001 (N=62)	-0.021 (N=52)

This table shows correlation coefficients for annual coal prices with annual natural gas prices and annual residential construction. The coal prices are the Energy Information Administration coal price series from the Annual Energy Review 2011, which extends from 1949 - 2010. The natural gas data are Energy Information Administration wellhead price data, which are available for 1922 - 2011. The construction data are the “New Privately-owned Units Started” series from the Census Bureau, which are available for 1959 - 2010. The observation numbers in the table are the number of years in which data are available for coal and the variable of interest.

Table C.2: Difference in differences specification test

Dependent variable:  $\ln(\text{wood chips per square mile})$ 

Treatment*1997	0.24	(0.28)
Treatment*1998	-0.29	(0.22)
Treatment*1999	-0.14	(0.23)
Treatment*2000	-0.24	(0.37)
Treatment*2001	0.28	(0.83)
Treatment*2002	0.03	(0.33)
Treatment*2003	0.08	(0.60)
Treatment*2004	-0.09	(0.48)
Constant	3.40	(0.13)***
Year fixed effects	Yes	
County fixed effects	Yes	
Observations	143	

For years before the PTC, this table shows a regression of the log of wood chip quantity on year\*treatment dummies, year dummies, and county dummies. Standard errors are clustered at the county level. The excluded year is 1996.

Table C.3: Robustness to different ways of dealing with missing price observations

	Drop missing	Min year
Supply elasticity (both instr.)	0.43 (0.3)	0.41 (0.2)*
Supply elasticity (single instr.)	0.60 (0.3)**	0.61 (0.2)***
First stage coeff. for coal price	2.40 (0.2)***	2.77 (0.2)***
First stage coeff. for PTC*after	0.2 (0.1)**	0.16 (0.1)
N	204	239

This table shows how the main result for the price elasticity of wood fuel supply is robust to different ways to dealing with missing observations on price. “Drop missing” removes observations with missing values. “Min year” replaces missing observations with the lowest value observed in that year regardless of PTC treatment status. Legend for significance stars: \* 10%; \*\* 5%; \*\*\* 1%.

Table C.4: Alternative time trend specification

Dependent variable	First stage Wood chip price	2SLS Wood chip quantity
Coal price	0.18 (0.02)***	
Wood chip price		26.1 (6.2)***
Linear time trends for treatment and control	Yes	Yes
County fixed effects	Yes	Yes
N	239	239

This table shows the results from estimation in levels with a linear time trend. The 2SLS model uses both coal price and PTC\*after as instruments. Standard errors in parentheses are clustered at the county level. Legend for significance stars: \* 10%; \*\* 5%; \*\*\* 1%.

Table C.5: Sensitivity analysis for area from which power plants source wood chips

	2 miles	25 miles	40 miles
Additional treated counties	Kennebec, Oxford	Somerset	Waldo, Hancock
Effect of PTC program	-0.01 (0.2)	0.04 (0.2)	0.15 (0.2)

This table shows how the list of “treated” counties and the estimate of the effect of the PTC change as the area from which power plants are assumed to source wood increases. Distances on tops of columns are radii from the power plant. The counties listed in each column are counties that do not contain biomass power plants, but that are added to the treatment group as this distance grows. The numbers for “Effect of PTC program” are the coefficients on PTC\*after in equation (3.3), including county fixed effects, all covariates, and exponential time trends for treated and control counties. In all cases, the program does not have a statistically significant effect on wood chip quantity.

Table C.6: IV results including pulpwood and sawlogs as covariates

Dependent variable:  $\ln(\text{wood chips per square mile})$

$\ln(\text{Price})$	0.64	(0.25)**	0.48	(0.26)*
$\ln(\text{Sawtimber per square mile})$	0.02	(0.16)	0.03	(0.16)
$\ln(\text{Pulpwood per square mile})$			0.22	(0.15)
$\ln(\text{Oil price})$	0.03	(0.15)	0.08	(0.14)
Exponential time trends	Yes		Yes	
County Fixed Effects	Yes		Yes	
Counties	16		16	
Observations	239		239	

This table reports 2SLS regressions of the log of wood chip quantity on wood chip prices and covariates, using coal prices as the instrument. Standard errors are clustered at the county level. Legend for significance stars: \* 10%; \*\* 5%; \*\*\* 1%.