Title
Essays in Energy and Environmental Economics

Author
Hausman, Catherine Helena

Publication Date
2013

Peer reviewed|Thesis/dissertation

Essays in Energy and Environmental Economics

by

Catherine Helena Hausman

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 Maximilian Auffhammer, Co-Chair
Professor Severin Borenstein, Co-Chair
Professor Peter Berck
Professor Lucas Davis

Spring 2013

Essays in Energy and Environmental Economics

Abstract

Essays in Energy and Environmental Economics

by

Catherine Helena Hausman
Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Maximilian Auffhammer, Co-Chair
Professor Severin Borenstein, Co-Chair

Energy production is associated with a number of significant environmental externalities. For instance, coal-fired power plants emit both local pollutants (such as particulate matter and sulphur dioxide) and the global pollutant carbon dioxide. This creates a need for government intervention: left to their own devices, energy producers will do more environmental damage than is socially optimal. The choices faced by policy-makers in regulating the energy industry are, however, rarely clear. Government regulators must trade off the externalities caused by different types of energy production. While nuclear power generation does not emit carbon dioxide, there is the risk of significant environmental damage in the event of a nuclear meltdown. While many proponents of biofuels hoped that replacing fossil fuels with biofuels would decrease carbon dioxide emissions, land-use changes associated with biofuels production can cause environmental damage. These trade-offs motivate the chapters of this dissertation.

In the first chapter, I study changes to nuclear power safety following major regulatory changes in electricity markets. Following electricity market restructuring, approximately half of all commercial U.S. nuclear power reactors were sold by price-regulated public utilities to independent power producers. At the time of the sales, some policy-makers raised concerns that these corporations would ignore safety. Others claimed that the sales would bring improved reactor management, with positive effects on safety. Using data on various safety measures and a difference-in-difference estimation strategy, I find that safety improved following ownership transfers and the removal of price regulations. Generation increased, and this does not appear to have come at the cost of public safety.

This paper contributes to several strands of the energy literature. First, it fits in with the literature on electricity deregulation. While this literature has considered a broad set of outcomes, my paper is the first to look closely at safety, an outcome of particular interest for nuclear energy. In line with that, it also contributes to the literature on nuclear safety, which has been of particular interest given accidents at Three Mile Island, Chernobyl, and Fukushima. Finally, my work is germane to the literature on the consequences of deregulation for outcomes beyond private efficiency gains. While there is now some consensus that deregulation can lead to the alignment of private costs and thus to efficiency gains, less is

known about the effect on external costs. Papers in this literature are necessarily industry-specific: the interaction of private cost reductions with changes to quality or changes to external costs is highly context-dependent. However, this paper provides intuition for the mechanisms at work, some of which are generalizable beyond the nuclear power industry.

In the second and third chapters, I study land-use changes relating to biofuels production. Transportation in the U.S. accounts for a significant portion of greenhouse gas emissions. Motor gasoline, excluding ethanol, accounts for around 20 percent of U.S. greenhouse gas emissions, or over 1 billion metric tons each year.[1] Biofuels have been promoted as an alternative to petroleum products that bypasses some of the fundamental problems with the oil market: supporters claim that it is renewable (whereas conventional oil is exhaustible), produced in the U.S. (as opposed to regimes in some cases unfriendly to the U.S.), and carbon-friendly. As the acreage devoted to biofuels crop production expands, however, it can compete with cropland used for food or with natural ecosystems.

In the second chapter, joint with Maximilian Auffhammer and Peter Berck, I examine price impacts of biofuels production. The last ten years have seen tremendous expansion in biofuels production, particularly in corn ethanol in the United States, at the same time that commodity prices (e.g., corn) have experienced significant spikes. While supporters claim that biofuels are renewable and carbon-friendly, concerns have been raised about their impacts on land use and food prices. This paper analyzes how U.S. crop prices have responded to shocks in acreage supply; these shocks can be thought of as a shock to the residual supply of corn for food. Using a structural vector auto-regression framework, we examine shocks to a crop's own acreage and to total cropland. This allows us to estimate the effect of dedicating cropland or non-crop farmlands to biofuels feedstock production. A negative shock in own acreage leads to an increase in price for soybeans and corn. Our calculations show that increased corn ethanol production during the boom production year 2006/2007 explains approximately 27 percent of the experienced corn price rise.

In the final chapter, I study land-use change in Brazil arising from biofuels production. Scientists and economists are increasingly worried that biofuels production is leading to deforestation, and hence loss of habitats and increased carbon dioxide emissions. I estimate land use changes in response to shocks in sugarcane (a biofuels feedstock) and soybean (thought to be affected by United States corn ethanol production) prices in Brazil at a national and regional level. Using county-level data from 1973 to 2005, I consider a dynamic panel data model of input demand for land, conditioning on price changes of other commodities. Unlike the existing literature, I apply a dynamic panel data estimator that is unbiased (unlike OLS with fixed effects) and more precise than GMM. The short-run price elasticity of sugarcane acreage in Brazil is estimated to be approximately zero, whereas the elasticity of soybean acreage is 0.9 when both spot and futures prices change. The regional estimates for soybeans show considerable variation, and are highest in areas of ecological importance, such as the cerrado. Sugarcane estimates are more homogeneous. These results should be taken into account in impact assessments of biofuels.

[1] Source: Energy Information Administration, Monthly Energy Review, April 2013.

To Josh and Alex.

Contents

Acknowledgments

I can think of no place I would rather have completed my graduate studies than at the University of California, Berkeley. Professors Severin Borenstein and Maximilian Auffhammer, my committee co-chairs, have inspired me with their intellectual curiosity, rigor, and honesty. Their support and advice have been invaluable. Professor Lucas Davis changed the way I think about research, and I thank him for that and for his good humor. Professor Peter Berck shepherded me (and most of my class) from day one of graduate school and somehow always knew when I needed a kick in the pants versus a word of encouragement.

I will greatly miss my intellectual home at UC Berkeley, the University of California Energy Institute. UCEI brings together great researchers who also happen to be great people. Thank you to Bev Alexander, Carl Blumstein, Casey Hennig, Catherine Wolfram, Erica Myers, Howard Chong, Josh Blonz, Judd Boomhower, Karen Notsund, Koichiro Ito, Liz Bailey, Matt Zaragoza, Meredith Fowlie, Patrick Baylis, Walter Graf, and all the visitors for friendships and engaging conversations.

Thanks to my classmates in Agricultural and Resource Economics: Anna, Alex, Chantal, Charles, Di, Gianmarco, Fei, Jessica, Lunyu, and Maya, for keeping things in perspective.

Finally, and most importantly, I can hardly begin to express all my thanks to my family. To my husband Josh, who also happens to be my favorite economist, I wouldn't be who I am without you. Thank goodness I came to Berkeley, so I could meet you. To my father Paul, for suggesting I study economics. To my mother Sharon, for thinking everything I do is wonderful. To my sisters Amy and Mindy for laughs and diversions. To Josh's family, Dan, Cathy, and David, for smart questions. I am so grateful to have you all for my family.

Chapter 1

Corporate Incentives and Nuclear Safety[1]

1.1 Introduction

In the past two decades, a dramatic change to the nuclear power industry has taken place: approximately half of all U.S. nuclear power plants have been sold off by price-regulated utilities and now operate in competitive markets. Surprisingly, there is little evidence on how ownership transfers have affected safety. This paper provides the first comprehensive analysis of the impact of these nuclear power plant divestitures[2] on safety. Using data on a variety of safety measures and a difference-in-difference estimation strategy, I find no evidence that safety deteriorated; for some measures, it even improved following divestiture. Moreover, for given levels of generation, safety substantially improved. Ownership transfers led to the alignment of private incentives to increase operating efficiency, and these gains do not appear to have come at the cost of public safety.

The deregulation of electricity generation markets, begun in the late 1990s, was undertaken in part to increase efficiency and lower costs. It was thought that, under rate of return regulation, incentives were not aligned for utilities to minimize costs in the generation portion of their business. Robust empirical evidence now shows that efficiency gains were indeed realized at both fossil-fuel-fired plants and nuclear plants after the restructuring of electricity markets. Davis and Wolfram (2012) attribute a 10 percentage point increase in operating

---

[1] I am grateful to Max Auffhammer, Peter Berck, Severin Borenstein, Lucas Davis, and Catherine Wolfram for their invaluable advice. I thank Michael Anderson, Meredith Fowlie, Daniel Hausman, Joshua Hausman, Koichiro Ito, Per Peterson, Charles Seguin, Anna Spurlock, and seminar participants at Calgary University, Dartmouth College, George Washington University, Rice University, the University of British Columbia, University of Colorado, University of California Berkeley, University of California Irvine, University of Michigan, Wellesley College, and Williams College for excellent comments. This work was supported in part under a research contract from the California Energy Commission to the Energy Institute at Haas. All errors are mine.

[2] As described below, divestiture refers to the process whereby utilities transfer generation assets to unregulated companies, and it can involve either transfer to an unregulated subsidiary of the regulated utility or sale to an independent power producer.

efficiency at nuclear power plants to divestiture from investor-owned utilities.

While there is now some consensus that electricity market restructuring led to the alignment of *private* costs and thus to efficiency gains, less is known about the effect of the market changes on *external* costs. Even as deregulation began in the late 1990s, some feared that the independent power producers purchasing nuclear plants would ignore safety concerns in the interest of maximizing profits. Others claimed that deregulation and consolidation would improve reactor management, and that the new owners would work hard to avoid costly plant shutdowns. David Lochbaum of the Union of Concerned Scientists was quoted in the *New York Times* as saying "[t]he new owner of a nuclear power plant clearly has a commitment to a nuclear future... you can also make the counterargument that the new owner is only trying to make a quick buck, to recoup their investment and make some money."[3]

My empirical strategy exploits the fact that only half of the reactors in the U.S. were divested and that the timing of divestiture varied widely. These differences in divestiture were largely the outcome of differential electricity deregulation legislation across states. I make the identifying assumption that this timing is exogenous to nuclear safety. To examine the validity of this assumption, I test for the possibility of selection bias. Looking at pre-divestiture safety records, I find no statistically or economically significant differences between the plants that later divest and those that remain controlled by investor-owned utilities.

Unfortunately, while catastrophic events may represent the largest social cost of nuclear power, their risk is not observable directly. I am, however, able to analyze data from the Nuclear Regulatory Commission (NRC) on five safety measures: initiating events (unplanned power changes), fires, escalated enforcement actions,[4] collective worker radiation exposure, and average worker radiation exposure. The NRC compiles these data from both operator reports and regular inspections. I choose these five measures in part because they may be the least open to manipulation by plant operators. Unplanned power changes, for instance, are not possible to hide from safety inspectors, since generation to the electrical grid is metered. Additionally, these measures represent a broad portion of the risk to plants. Initiating events cover a large portion of the internal event core damage risk to nuclear plants (Eide, Rasmuson, and Atwood 2005). Also, the NRC's authority to use escalated enforcement actions "extends to any area of licensed activity that affects the public health and safety;"[5] I thus use these as the best measure available of the failure of a reactor's operator to follow federal safety regulations.

I find that divestiture leads to a 17 percent reduction in the expected number of initiating events, a 46 percent reduction in the expected number of fires, and a 35 percent reduction in the expected number of escalated enforcement actions. While the point estimates are not very precisely estimated, the magnitude of the coefficients is economically significant. Furthermore, moderate increases in the number of events can be ruled out at the five percent

[3] Wald, Matthew L. 2000. "Safety a Worry as Companies Shop for Nuclear Reactors" *New York Times*, February 22.

[4] As described in the data section, escalated enforcement occurs when the Nuclear Regulatory Commission imposes notices of violation and/or financial penalties on plants it deems out of compliance with safety regulations.

[5] http://www.nrc.gov/about-nrc/regulatory/enforcement/program-overview.html (Accessed July, 2011).

level. The results are robust to a number of specification checks, including various count models and OLS estimation. For radiation exposure, I find a reduction of 25 percent for collective worker exposure and 18 percent for average worker exposure. I also examine the effect of divestiture on safety for given levels of generation. This is important because the results described above include an indirect generation effect. The direct effect of divestiture on unsafe events is negative, but divestiture also increases generation, thereby increasing the exposure of the plant to an event. I find larger reductions in the expected number of unsafe events for given levels of generation, and the results are statistically significant at the 1 percent level for initiating events and 5 percent level for escalated enforcement.

The results are stable across reactor type and location, alleviating concerns about selection bias. In specifications allowing for differential trends, I find that divested plants and non-divested plants were on similar trends prior to treatment. After plants are divested, they improve over time relative to non-divested plants. These results are reassuring that the difference in safety records is not driven by temporary changes immediately following divestiture.

These findings are consistent with the incentives faced by nuclear plant operators, who have strong incentives to avoid outages. Because wholesale electricity prices are much higher than variable costs for nuclear plants, any outage leads to large losses in operating profits. Thus unsafe events that lead to plant shutdowns incur private costs for plants beyond the costs of the repairs themselves. On the other hand, maintenance to prevent unsafe events is also costly if it requires a plant to shut down. Prior to divestiture, plants may have been able to pass on some of the costs of outages to their ratepayers; since this is not possible in competitive generation markets, divestiture likely changed their incentives for maintenance. Ex-ante predictions about the effect of divestiture on maintenance are not possible, for reasons discussed below. However, both anecdotal evidence and the empirical results suggest that divestiture led to improved plant management and thus to better safety records.

This paper contributes to several strands of literature. First, restructuring transformed the electricity industry in many parts of the U.S., stimulating interest among economists and policy makers in understanding the consequences of these broad market reforms. This literature is part of a larger literature on the evolution of markets following deregulation. Electricity serves as a useful empirical setting in this broader literature for a few reasons: (1) electricity is a homogeneous good, so quality changes do not confound the analysis; (2) some states deregulated while others did not, and the timing of deregulation varied. This process, while not random, has generally been thought to be exogenous to power plant operations. Several important outcomes have been analyzed in this context, including operating efficiency (Bushnell and Wolfram 2005; Davis and Wolfram 2012; Fabrizio, Rose, and Wolfram 2007; and Zhang 2007), market power (Borenstein, Bushnell, and Wolak 2002; Bushnell, Mansur, and Saravia 2008), and emissions (Fowlie 2010). This paper is the first to analyze safety, which plays a crucial role in energy production and particularly in nuclear power. Nuclear power is controversial precisely because of the potential for catastrophic events, so understanding how deregulation impacted the probability of unsafe events is crucial.

This paper also contributes to the literature on nuclear power safety. Analyses of nuclear power safety emerged following accidents at Three Mile Island and Chernobyl (e.g. David, Maude-Griffin, and Rothwell 1996, Feinstein 1989, Hanemann et al. 1992, Rothwell 1989, and Rust and Rothwell 1995), and the recent accident at the Fukushima Daiichi facility has renewed interest in understanding the risks the public faces from nuclear plants. This paper does not claim to answer the broad questions of whether the world should use nuclear power to meet its energy needs or of how safety should be regulated. It does, however, speak to how a major market transformation in the U.S. impacted almost half of the nuclear fleet.[6] Moreover, it relates to the wider literature on the structure of the nuclear power sector (including Davis 2012, MIT 2003, and MIT 2009). This sector comprises a significant portion of the U.S. electricity industry, and interest in it has been renewed in recent years because of its status as a low-carbon source of large-scale baseload electricity generation.

Third, this paper is germane to the literature on the consequences of deregulation for outcomes beyond private efficiency gains. When the airline industry was deregulated, for instance, concerns were raised about airline safety (Barnett and Higgins 1989, Golbe 1986, Kennet 1993, and Rose 1990). Importantly, though, one of the main mechanisms through which safety and profitability are related in air travel is in the consumer's demand function; this mechanism is not expected to operate in the case of nuclear power generation, as electricity is not differentiable for end-users. In related work, water privatization led to concerns about increases in water-borne illness (Galiani et al. 2005). Papers in this literature are necessarily industry-specific: the interaction of private cost reductions with changes to quality or changes to external costs is highly context-dependent. However, this paper provides intuition for the mechanisms at work, some of which are generalizable beyond the nuclear power industry.

1.2 Background and Related Literature

1.2.1 Electricity Deregulation

Deregulation refers to the broad set of reforms proposed for the U.S. electricity sector in the late 1990s; the set of reforms actually implemented and their timeline varied by state. Prior to deregulation, and in states where deregulation did not occur, local monopoly utilities bundled generation, transmission, and distribution services. Local public utilities commissions (PUCs) set the prices the utilities received so the utilities could recover fixed costs plus a fair rate-of-return; one example of such regulation is average-cost pricing. This cost of service pricing is the most extreme form regulation took; typically, some incentives for generators to keep costs low were built into the regulatory process. During deregulation, proposed reforms included separating generation, transmission, distribution, and retailing components of the sector and applying various reforms to each of these. Generation was opened to competition (with transmission and distribution still considered natural monopolies), and prices, entry and exit were deregulated. Retail reforms allowed consumers to choose between competing suppliers. Overviews of the economic and political arguments

[6] One related paper is Verma, Mitnick and Marcus (1999), which finds mixed results for the effect of incentive regulation programs prior to divestiture on power plant safety.

motivating electricity deregulation, the various forms deregulation could take, and the ex-ante concerns about deregulation can be found in Joskow (1997) and White (1996). As of 2010, fifteen states and the District of Columbia had restructured their electricity sector.

Divestiture refers to the process whereby utilities transfer generation assets to unregulated companies. This can refer to either transfer to an unregulated subsidiary of the regulated utility or sale to an independent power producer. In some states, this was required by legislation, to prevent market power following deregulation. For nuclear power reactors, this entry into competitive wholesale markets is the main component of deregulation expected to affect operations.

The main economic argument for generation deregulation was to increase efficiency and lower costs. Efficiency gains with deregulation are generally thought to come from aligning incentives vis-a-vis input choices, as in the Averch and Johnson (1962) model or from correcting agency problems, as in the Laffont and Tirole (1986) model. For overviews of these models and their extensions, see Baron (1989) and Kahn (1988). There is robust empirical evidence of efficiency gains at power plants in the U.S. following deregulation (Fabrizio, Rose, and Wolfram 2007 and Davis and Wolfram 2012).

An important assumption of this paper is that electricity deregulation was exogenous to nuclear power plant performance. The rationale for this assumption is that divestiture was tied very closely to state-level electricity deregulation, which was driven by a host of political and economic factors (Ando and Palmer 1998). Past nuclear power plant construction certainly was one motivator for deregulation, through the "stranded costs" problem. Since electricity prices were set at average rather than marginal cost, historical nuclear construction led to regulated electricity rates that were much higher than wholesale prices. Thus states with high historical nuclear fixed costs may have been more likely to deregulate (Griffin and Puller 2005, Joskow 1997, and White 1996). Davis and Wolfram (2012) find a slightly higher construction cost for plants that were eventually divested, however the difference is small (4 percent) and not statistically significant. Any difference in past nuclear construction costs should be time-invariant, and as such can be controlled for in empirical specifications with fixed effects. Finally, to my knowledge, poor nuclear safety records did not play a role in electricity restructuring.

Note that the impact of divestiture should be interpreted as including three endogenous features. First, it is possible that a utility seeking to sell its nuclear reactor would invest in plant improvements prior to the sale. This is particularly likely at poor performers, which utilities might be afraid they would be unable to sell. Second, while the act of divestiture may be exogenous to plant characteristics and performance, which company buys the plant is not exogenous. That is, there were several companies that purchased divested reactors, and they likely sorted on plant characteristics. Neither feature of deregulation affects the validity of the empirical estimation in this paper, but rather the mechanisms through which the impact of divestiture operates. Additionally, the timing of divestitures following deregulation may be endogenous. This is examined in the empirical analysis that follows.

1.2.2 Nuclear Power

There are currently 65 nuclear power plants in the U.S., accounting for 10 percent of total electric capacity. Because nuclear power plants are "baseload," meaning that they run around the clock, they contribute 20 percent of total electricity generation (NRC 2010). Most of the nuclear plants in the U.S. have multiple reactors, and there are currently 104 operating reactors. There are two types of reactors in the United States, pressurized-water reactors (PWRs) and boiling-water reactors (BWRs). In both types of reactor, fuel assemblies containing enriched uranium create heat, which then produces steam to turn a turbine.

Nuclear power plants have both advantages and disadvantages relative to fossil fuels plants. Once a nuclear power plant is built, its marginal costs are low. Furthermore, it emits no carbon dioxide during operation. Nuclear power also has advantages over alternative energies such as wind and solar, as it is not intermittent. Also, it can theoretically be built in areas where wind and solar are cost ineffective and hydroelectric resources are unavailable. However, nuclear power has several large disadvantages. Plants are expensive to build, so the levelized cost of nuclear power may be higher than that of fossil fuel plants (Davis 2012). Accidents at nuclear power plants can be catastrophic, and the public has been understandably wary in the wake of the events at Three Mile Island (in 1979), Chernobyl (in 1986), and Fukushima (in 2011). An additional concern is the potential for terrorists to acquire radioactive materials or attack U.S. nuclear sites. Finally, one of the main issues raised by environmentalists is the treatment, storage, and transport of spent nuclear fuel. Spent fuel assemblies can be stored in pools or dry casks at power plants. As of 2009, approximately 60,000 metric tons of spent fuel were stored at power plants (NRC 2010).

Nuclear power plant safety is regulated in the U.S. by the Nuclear Regulatory Commission (NRC), a government agency. The NRC also regulates nuclear research facilities and radioactive waste. It is responsible for licensing and inspections. The NRC has the ability to require unsafe plants to shut down; it can also apply fines for safety violations. The NRC does not appear to enforce its safety regulations differentially between price-regulated and divested plants.[7] In addition to the government oversight by the NRC, nuclear power reactor safety oversight is carried out by the Institute of Nuclear Power Operations (INPO). INPO is an industry organization that conducts reactor inspections and facilitates best-practices sharing across operators. It was founded following Three Mile Island, as operators realized an incident at any one plant had the potential to lead to the closure of all plants (Rees 1994).

Finally, incentives for safety are affected by liability insurance, which is highly regulated. Both investor-owned utilities and independent power producers are regulated according to the Price-Anderson Act (PAA). The PAA has a three-tiered liability system for all facilities. Nuclear power operating companies are required to purchase the maximum insurance coverage available in the private market, \$375 million annually as of 2010. The second tier is a joint pool; companies are required to pay retrospective premiums in the event of an accident. Companies must prove to the NRC that they will be able to make these payments by, for instance, posting a bond. Retrospective payment is currently set at approximately

[7] See, e.g., the NRC's policy statement in the Federal Register regarding electricity market restructuring: "Final Policy Statement on the Restructuring and Economic Deregulation of the Electric Utility Industry." *Federal Register* 62:160 (19 August 1997): 44071-44078.

$112 million per reactor per incident. The federal government is responsible for all payments above this primary and secondary coverage. The Price-Anderson Act covers liability claims but not on-site damages; the NRC separately requires companies to maintain funds for these damages.

1.2.3 Incentives and Safety

To provide context for the empirical results that follow, I discuss the incentives for safety faced by nuclear power plants; a formal derivation of this model is given in the following section. Most of the costs of a nuclear power plant are fixed and are incurred at the time of construction. The plants' marginal cost is much lower than the market price of electricity generation, which is determined by the marginal cost of the marginal plant.[8] According to a recent Energy Information Administration report (EIA 2011), variable costs are 2.17 cents per kilowatt-hour for nuclear plants and 4.05 for fossil-fuel steam plants. As such, even when demand is very low, nuclear plants can earn large operating profits. Thus they generally run continuously except for outages related to repairs and refueling.[9] Any outages, planned or unplanned, lead to large losses of operating profits.[10] Maintenance decisions are thus, in part, a trade-off between incurring downtime for plant repairs and preventing unplanned outages. This trade-off is less relevant if maintenance can be conducted while the plant is still generating.

Consider a profit-maximizing plant choosing a level of maintenance, which affects either reliability (i.e., avoiding unplanned outages), safety, or both. The costs of unreliability are private (limited to lost revenues for the plant) while safety costs are social (representing risk to the general public).[11] In many cases, the maintenance that reduces outages has complementarities with safety (MIT 2003). The firm chooses the level of maintenance that equates private marginal benefits (e.g., avoiding unplanned outages) with private marginal costs (maintenance costs as well as foregone revenues if the plant must be down for repairs). Since not all safety costs are internalized, the firm chooses a lower level of safety maintenance than is socially optimal. Furthermore, if the same maintenance improves both reliability and safety, the sub-optimal level of maintenance leads to socially sub-optimal levels of both reliability and safety. If, however, reliability maintenance and safety maintenance are unrelated, the firm will choose the socially optimal level of reliability but a sub-optimal level of safety.[12]

[8] For representative supply and demand curves showing nuclear marginal costs compared to fossil fuel costs, see Mansur (2008) or Griffin and Puller (2005).

[9] Vary rarely, nuclear plants are asked to reduce generation to preserve stability on the electrical grid.

[10] There is potential for the owner of a nuclear power plant to use outages to exercise market power, if it owns other generators. However, if the other generators have higher marginal costs than the nuclear plant, exercising market power by shutting down the nuclear plant is not the first-best strategy of the firm. Rather, the firm would take the higher cost plant offline. Moreover, if the nuclear power plant has a firm contract to sell, the owner will be required to purchase replacement power when the plant is down. Since the replacement power is more costly than the nuclear plant's generation, the firm has no incentive to exercise market power by taking the nuclear plant offline.

[11] Indeed, liability for nuclear power plants is capped in the U.S. under the Price-Anderson Act described in section 1.2.2.

[12] It should be noted that the Nuclear Regulatory Commission sets standards on safety-related equipment;

The incentives are less clear under rate of return regulation. As described in section 1.2.1, prices in regulated electricity markets are set so that monopoly utilities recover their costs. Variable costs are passed on to rate payers, and utilities are additionally allowed a fair rate-of-return on their fixed costs. If the regulatory compact is that the utilities commission will allow the utility to pass on all costs to consumers, then the regulated plant has no incentive to minimize costs. In practice regulation usually involves some incentives for generators, but utilities are typically able to pass on to consumers a greater portion of their costs than are independent power producers.

In regulated environments, when a nuclear power plant is not generating, the utility will substitute with a more expensive plant (for instance, natural gas-fired), and then pass on this higher generation cost to its customers. Thus the incentives to avoid unplanned outages may be lower at a plant operating under rate-of-return regulation. In the short term, this is mitigated by the ability of the regulated plant to pass on its maintenance costs. In the long term, however, a deregulated plant may have a greater incentive to improve technical efficiency to lower maintenance costs.

As described in section1.2.1, a key argument for electricity deregulation was to increase efficiency by aligning cost incentives and correcting agency problems. Davis and Wolfram (2012) find that reactors are available to generate for a significantly higher percentage of the time following divestiture. This improved operating efficiency appears to have come in the form of shorter refueling outages, enabled by changes in management practices. One newspaper article describes Entergy, one of the larger owners of divested plants, flying a specialist and his equipment on the company jet from one reactor to another to fix an electrical generator.[13]

Where practices that improve reliability also improve safety, divested plants may have similarly improved safety records. For instance, unplanned outages and power changes, which represent both reliability and safety costs, might be expected to fall following divestiture. This possibility is explored in the empirical section of this paper. On the other hand, safety incidents that do not affect plant reliability may not fall after divestiture. In the case where safety and reliability are uncorrelated, the effect of divestiture on safety will depend on whether the divested plant internalizes more or less of the cost of a safety event. Liability under the Price-Anderson Act does not differentiate between plants owned by investor-owned utilities and those owned by independent power producers. However, divested plants could internalize more or less of the cost of a safety event if, for instance, they are subject to a differential level of public scrutiny following an accident or place differential value on reputation.[14]

As described in section 1.2.2, the Nuclear Regulatory Commission regulates nuclear safety in the U.S., and INPO is an industry self-regulation organization. With perfect information

---

if the equipment malfunctions, the plant must shut down. Any maintenance on equipment the NRC observes and regulates is therefore related to generation.

[13] Wald, Matthew L. 2001. "Despite Fear, Deregulation Leaves Nuclear Reactors Working Harder, Longer and Safer." *New York Times*, February 18.

[14] The empirical portion of this paper examines whether consolidation affected safety records: companies that own many plants may internalize more safety costs if an incident at one plant leads to scrutiny at all plants.

and regulatory oversight, the socially optimal level of safety could be achieved in both the price-regulated and competitive generation markets. Note that there is still room for safety to improve following divestiture: if divested plants attain greater technical efficiency because of the alignment of cost incentives for reliability, the socially optimal level of maintenance would be higher.

In addition to the framework I present here, two theoretical models could be applied under price regulation: (1) the Averch-Johnson (1962) model, in which firms over-invest in capital, and (2) agency models, such as Laffont and Tirole (1986), in which firms exert sub-optimal levels of effort. Averch and Johnson show that plants under rate-of-return regulation over-invest in capital relative to labor. The intuition is simple; under rate-of-return regulation, a firm's profits are a function of its capital investments. If the allowed rate-of-return on investment is higher than the firm's cost of capital, the firm over-invests in capital relative to labor. The Averch-Johnson effect may explain the construction of nuclear power plants, but it is likely not relevant in the operation of nuclear plants. A long history of cost overruns in nuclear power plant construction meant that many local regulators were wary of approving further capital expenditures (Joskow and Schmalensee 1986).

Fabrizio et al. (2007) cite agency models in explaining why deregulation may improve operating efficiency at thermal power plants. In agency models such as Laffont and Tirole (1986), efforts to run a firm efficiently by reducing costs provides some disutility to the firm's manager. The regulator fails to compensate the manager for this disutility, perhaps because effort is unobservable or unverifiable, so the manager exerts less effort than is socially optimal. For nuclear plants, efforts to maintain reliability and safety are unobservable to public utilities commissions, since outages and accidents are stochastic. A manager could exert minimal effort while blaming outages and accidents on bad luck. In the case of nuclear plants this is likely mitigated by an aversion on the part of both the manager and the public utilities commission to the public scrutiny that follows extended outages or severe accidents. In that case, managers would be more willing to exert effort to maintain safety and reliability, and regulators would be less willing to treat outages and accidents as bad luck.

Overall, the impact of deregulation and divestiture on plant safety is theoretically ambiguous. It depends crucially on a number of issues, including (1) whether state regulators allowed the monopoly utility prior to divestiture to pass on maintenance costs and/or replacement power costs; (2) whether maintenance for reliability has additional safety benefits; (3) whether divested plants internalize more or less of the cost of safety events; and (4) the level of federal[15] safety regulations. Since many of these factors are unobservable, I next turn to empirical evidence.

[15]All safety regulations are administered at the federal level. However, public scrutiny may vary across states.

1.3 Data

1.3.1 Power Plant Safety

For the empirical analysis, I compile data on a variety of risks to nuclear plants. The Nuclear Regulatory Commission (NRC) tracks a number of safety measures for all reactors in the United States. Reactor operators are required, under the Code of Federal Regulations (10 C.F.R. §50), to provide reports to the NRC following any shutdown, deviation from technical specifications, or event resulting in degraded plant safety. These licensee event reports contain information by date and by plant on the specific event or condition involved, including narrative descriptions. These are publicly available from the NRC. Additionally, the NRC performs regular plant inspections. These can involve inspectors permanently stationed at the plant, regional inspectors, and inspectors for specific areas such as on-site security. Inspections may involve reviewing records, observing drills and simulations, observing maintenance procedures, and testing equipment. Results are made public by the NRC.

The NRC additionally synthesizes and publishes data on safety measures of particular interest for this study:

- initiating events, including unplanned outages and power changes
- fires
- worker radiation exposure
- escalated enforcement actions, including orders and fines

Data are available since 1988 on initiating events in the report "Rates of Initiating Events at U.S. Nuclear Power Plants 1988–2010." All scrams (or trips), which are unplanned outages, are categorized as initiating events. Unplanned power changes that are not scrams are also categorized as initiating events. Each initiating event is assigned to one of several categories, such as "stuck open safety relief valve" or "loss of feedwater." One advantage of analyzing initiating events is that they represent a significant portion of the known internal risk to plants (Eide, Rasmuson, and Atwood 2005). These reactors trips are frequently used as a summary measure of reactor safety. They indicate that some safety system was triggered, and the rapid power-down can itself subject the plant to additional risk (David, Maude-Griffin, and Rothwell 1996). Since initiating events correspond to unplanned loss of power (either total loss of power, as in a scram, or partial loss of power), these are events in which reliability maintenance overlaps with safety maintenance.

I also analyze fires, a safety event of particular interest to the NRC, for which I have data since 1990. The NRC dataset, "Fire Events Data from Licensee Event Reports," gives the original source document citation, the event date, the plant's mode at the time of the fire (e.g., power operating, refueling), operating capacity on the date of the fire, the physical area involved, and whether a safety alert was declared. Following an extensive fire at the Browns Ferry plant in 1975, the NRC revised fire regulations. The NRC now performs fire inspections on a regular basis and analyzes fire events for national trends. However, as recently as 2008, the Government Accountability Office (GAO) released a report calling for stricter regulations. The consequences of a fire depend on both where the fire starts and on how rapidly the fire can be extinguished. According to the GAO (2008) report, "[t]he most commonly reported cause of fires was electrical followed by maintenance-related causes and the ignition of oil-based lubricants or coolant. Although 13 fires were classified as significant alerts, and some of these fires damaged or destroyed unit equipment, NRC officials stated that none of these fires degraded units' safe shutdown capabilities or resulted in damage to nuclear units' core or containment buildings" (p 4). The report concluded that the NRC still needs to resolve several long-standing issues.

Additionally, I observe annual radiation exposure to individuals at the plant since 1974, using data from the NRC's "Occupational Radiation Exposure at Commercial Nuclear Power Reactors and Other Facilities (NUREG-0713)." Plants are required to report the radiation exposure of each monitored worker to the NRC, which reviews radiation control and monitoring during its regular plant inspections.[16] Monitoring procedures vary over time, but details of the regulation are given by 10 C.F.R. §20, "Standards for Protection against Radiation," which describes the "as low as (is) reasonably achievable" guidelines for radiation doses. Since the number of individuals could systematically vary across time (for instance, if divested plants employ fewer people), I analyze two separate measures. The first is collective worker radiation exposure, which sums exposure across all people; the second is average worker radiation exposure, which normalizes by the number of individuals monitored.[17] Data are at the annual facility level, in contrast to the other measures. Reporting procedures at plants with both operational reactors and permanently shut-down reactors vary: at some facilities, radiation exposure is reported separately for each reactor, whereas at some facilities they are reported in a combined measure. I drop observations that combine operational and permanently shut-down reactors.

A final measure of interest for safety is on "escalated enforcement," and is available in the form of the NRC dataset "Escalated Enforcement Actions Issued to Reactor Licensees." This tracks, since 1996, the notices of violation and penalties the NRC has imposed on reactors,[18] ranked according to severity level. It is part of the NRC's enforcement program, which focuses on compliance with regulatory requirements and identification and correction of violations. Currently, the NRC evaluates seven areas of safety: initiating events, mitigating systems, barrier integrity, emergency preparedness, occupational radiation safety, public radiation safety, and security. Three sanctions are possible: notices of violation (NOVs), civil penalties (i.e., monetary fines), and orders (e.g., to suspend operations). Minor violations are documented, but the lowest level of violations are not part of the "escalated enforcement" program. For each case, the NRC publicly posts the violation type (NOV and/or order) and

[16]For instance, a 2003 inspection report for Beaver Valley described NRC review of personnel dosimeters; frisking instruments; radiation portal monitors; protective clothing and self-contained breathing apparatus; radiological work permits; and daily health physics status meetings.

[17]The collective exposure measure, summing across workers, may be the most relevant measure of overall exposure. If, however, there are nonlinearities in the dose response function, then the average exposure for individual workers is also of interest.

[18]For plants with multiple reactors, notices of violation and penalties may refer to only one reactor, but more commonly refer to all the reactors at the plant.

severity, the amount of any civil penalty, the date issued, and a short description. This measure tends to lead to public scrutiny; the NRC may call a public meeting or issue a press release, and the violations are often reported by the media.

Unfortunately, the potential for catastrophic failure at a nuclear power reactor is not directly observable. I use these five measures because they are indicative of how well a plant is being maintained and how much risk the plant faces. As described above, initiating events represent a large portion of the known internal risk to plants and are widely used as a summary statistic of safety. Escalated enforcement represents the best available knowledge of the NRC about risk relating to a broad set of safety concerns. A second feature of the measures used is that safety along these dimensions is positively correlated with the plants' ability to generate electricity, matching the intuition described in section 1.2.3. However, for safety concerns that are not correlated with the ability of the plant to sell electricity, there is the possibility that divestiture will lead to increased risk. Two examples of these safety concerns might be maintenance of spent fuel storage and on-site security. However, it is important to note that if the NRC observes these actions and can shut down plants that violate regulations, this risk will also be correlated with the plants' ability to generate.

A third feature of the variables used in this analysis that is they represent measures it would be difficult for plant operators to hide or manipulate. However, to examine the possibility that divested plants are more likely to hide safety concerns, I have collected two additional measures. The NRC can initiate escalated enforcement procedures for violation of 10 C.F.R. §50.9, "Completeness and accuracy of information," if it determines that a plant operator withheld information. Escalated enforcement is also initiated for violation of 10 C.F.R. §50.7, "Employee protection," when plant operators discriminate against workers who raise safety concerns. These violations are infrequent, making empirical analysis difficult. However, as shown in Appendix C, I find that divestiture is associated with a lower rate of both types of violation, alleviating concerns about deregulated plants hiding safety concerns. Overall, the safety measures I use are thus indicative of the risk of catastrophic events. These measures may miss other types of catastrophic risk. However, for the measures used in this paper to not be informative of some other risk to plants, it would need to be the case that the risk was not positively correlated with my measures (e.g., required separate maintenance procedures), was not correlated with generation (so that the plants incentives were not aligned), and was either not observed or not enforced by the NRC.

1.3.2 Generation and Divestitures

Generation data, from Davis and Wolfram (2012), are published in the U.S. Department of Energy, Energy Information Administration (EIA) Power Plant Report (EIA-923). This survey (previously published as the EIA-906 and EIA-759 reports) provides monthly net generation in megawatt-hours for each nuclear reactor. I include only reactors operating as of January 1, 2000; this excludes a few reactors that were closed during the 1980s and 1990s.[19] To calculate capacity factor, I normalize generation by reactor design capacity.

[19] Most of these reactors were small and experimental. Exceptions include Browns Ferry 1, Millstone 1, and San Onofre 1.

Reactor design capacity is from the EIA "Nuclear Power Generation and Fuel Cycle Report 1997, Appendix C: Nuclear Units Ordered in the United States, 1953-1996." Divestiture dates, also from Davis and Wolfram (2012), are compiled from the EIA and cross-checked against SEC filings. For the empirical analysis that follows, I focus on these divestiture dates rather than deregulation dates. Since divestiture and deregulation are highly correlated for nuclear plants, I cannot separately identify the effect of regulation changes and ownership changes. I focus on changes in ownership, for which the timing can be more precisely defined than for changes in electricity market legislation; the related literature disagrees on which dates to use for electricity deregulation.[20]

1.3.3 Summary Statistics and Pre-Treatment Observables

Table 1 gives summary statistics on the five safety measures of interest plus generation and capacity factor for all 103 power reactors used in my analysis.[21] The average reactor has slightly fewer than one initiating event per year. Fires are quite rare. Worker radiation exposure averages 116 person-REMs per year. In 2008, this corresponded to roughly 1,300 workers per facility with an average dose of 0.1 rem; for comparison, the average person in the U.S. receives 0.3 rem from background sources of radiation and 0.3 rem from man-made sources (NCRP 2009). The average unit has one escalated enforcement intervention every two years, while producing over 7 million MWh of electricity. The average capacity factor was 88 percent. Note that capacity factor can be negative, since generation measured is net, rather than gross. It can also be greater than 100 percent, because of uprates that allow the unit to produce more generation than the initial design allowed.

To examine the potential for selection bias, table 2 shows mean values for each variable by the reactors' eventual divestiture status. Data are from 1996-1998; 1996 is the first year for which all safety measures are available, and 1998 is the last year in which no plants are divested. Observations are annual, and test statistics are adjusted for clustering at the plant level. Panel A shows that the safety measures are not statistically different at the 5 percent level between the plants that later divest and those that do not.[22] Panel B shows that reliability measures are statistically different at the 5 percent level; plants that were later divested have lower generation levels and capacity factors. As Davis and Wolfram (2012) discuss, reactors that were later divested had much lower generation in the late 1990s, which is explained by several long outages at a few plants.

Appendix C gives tests for differences in observable fixed reactor characteristics, previously analyzed in Davis and Wolfram (2012). There is a statistically significant difference in the proportion of boiling water reactors (BWRs) divested versus pressurized water reactors (PWRs). As such, I will test whether the effect of divestiture is robust to considering each type separately. There is no significant difference in age, capacity, number of reactors at each plant, or manufacturer (with the exception of reactors made by General Electric). There is

[20] See section 1.4.6 for a discussion of the timing of deregulation versus divestitures.

[21] There are currently 104 reactors in operation. For the empirical section of this paper, I drop Browns Ferry 1. This reactor was shut down from 1985 to 2007, and re-opened only following substantial investment.

[22] The regressions in section 1.4.8, "Dynamic Effects" also test for differences in pre-treatment trends.

a difference in the location of the divested facilities; this is not surprising, given the regional differences in deregulation patterns. To address concerns about selection bias, I later examine the robustness of the main results to excluding certain states and regions. Finally, no statistically significant difference is seen for maximum generating capacity, a measure that incorporates uprates and should be positively correlated with capital investment (Davis and Wolfram 2012). This further alleviates concerns about selection bias.

1.4 Empirical Evidence

1.4.1 Graphical Analysis

First, I plot an event study graph of the effect of divestiture on each safety measure at the quarterly level for all plants, intended to motivate the regressions that follow. This plot has the advantage of allowing me to examine pre-treatment trends in the number of unsafe events. While table 2 showed no difference in the pre-divestiture mean levels of unsafe events, this plot looks more flexibly at trends. Specifically, I plot the coefficients β_j from the following regression:

$$event_{i,t} = \sum_{j=-19}^{32} \beta_j \cdot 1[\tau_{i,t} = j]_{i,t} + v_t + \varepsilon_{i,t}$$

where $\tau_{i,t}$ denotes the quarter relative to divestiture, with $\tau_{i,-19}$ denoting nineteen quarters prior to divestiture, $\tau_{i,0}$ denoting the quarter of divestiture, etc.[23] The dummy variables v_t are quarter-of-sample effects. Thus the plotted coefficients β_j compare case reactors to the control reactors that never divest, net of time effects. The time effects play an important role, as unsafe events have generally been trending down; not including them would thus overstate the effect of divestiture. Figure 1 shows this for the sum of the three count variables: initiating events, fires, and escalated enforcement.[24] The figure additionally shows a lowess smoother in the pre-divestiture and post-divestiture periods in dashed grey lines. There is a decrease in incidents following divestiture, although it is smaller than the quarterly noise. The effect is not immediate, implying that there may be an adjustment period following divestiture, or there may be learning over time at divested units. The variance in the measure appears to decrease following divestiture; this is likely a direct implication of the count nature of the data. For a Poisson process, for instance, any reduction in the mean will also imply a reduction in the variance.

[23]The plot only shows event quarters for which there are at least 100 observations on divested units (approximately 70 percent of the full sample of divested units). Thus, while it is not a balanced panel, the sample does not change much.

[24]Summing across the three variables is an imperfect measure, since some double-counting is involved. For instance, a fire may set off an initiating event, or a severe initiating event may trigger escalated enforcement actions. As such, this measure is meant merely to serve as an illustration. The empirical analysis that follows considers each variable separately. Appendix B shows the plots for each individual type of event.

1.4.2 Regression Analysis

I next provide formal tests of the effect of divestiture on safety by regressing the safety measure on a divestiture dummy and a set of reactor fixed effects and year effects. For the three count variables (initiating events, fires, and escalated enforcement), the preferred specification is an unconditional negative binomial.[25] OLS is not expected to perform well given the count nature of these variables, although OLS results are shown along with other robustness checks. The negative binomial specification is preferred over a Poisson regression, which is subject to faulty inference if the data are overdispersed.[26] Poisson regression results are shown in the robustness checks. For specifications using radiation exposure, OLS regressions are used since radiation exposure is a continuous variable. These data are collected by plant rather than reactor, so I include facility fixed effects. For all specifications, standard errors are clustered at the plant level to allow for arbitrary correlation across reactors within a plant and across time.

One limitation of the estimates given in the previous two equations is that they are for the net effect of divestiture on safety, and are composed of two effects: the direct effect of divestiture plus an indirect effect through generation. That is, since the plants are operating for a greater percentage of the time, they may be more exposed to unsafe events. Hence an alternative outcome of interest is not the overall effect on safety, but rather the effect on the number of unsafe events for a given level of generation. One way to allow for this possibility empirically is to scale the safety variables by capacity factor (realized generation as a percent of design capacity) in each year; this is analogous to the engineering analyses that scale by reactor critical-years. This approach is not feasible at a monthly level; noise is introduced by large outliers in months when unsafe events occur despite very low capacity factors. These outliers can occur, for instance, if an unsafe event occurs early in the month and is then followed by an extended outage. Regressions at the annual level largely alleviate this problem; they smooth across months with low capacity factors. As such, all regressions are run at the annual level. For the results shown, I have dropped the approximately thirty observations for which capacity factor is less than 0.01.[27] For the count variables, the normalization is accomplished by including capacity factor as an exposure variable (i.e., as a regressor, with the coefficient on the logged variable equal to 1) in the negative binomial specification. For

[25]For an unconditional negative binomial specification, the individual effects α_i enter as dummy variables. This can leads to an incidental parameters problem for short panels, although simulations have found the resulting bias to be small (Allison and Waterman 2002). Conditional negative binomial specifications are not subject to an incidental parameters problem, however they have the unfortunate feature of allowing for heterogeneity across units only in the variance, and not in the mean. These specifications are shown in the robustness checks.

[26]The Poisson process assumes equality of the mean and variance, whereas in empirical settings the variance is often larger than the mean. This overdispersion leads to faulty inference (Type I error), with the null hypothesis rejected when it should not be. Fixed effects partially alleviate the problem by requiring only that the mean and variance be equal *within* groups, thus allowing for greater heterogeneity. Overdispersion tests, available upon request, indicate overdispersion for initiating events (with dispersion parameter approximately 0.2 to 0.3). They fail to reject equidispersion for fires, with dispersion parameter less than 0.01. The tests are inconclusive for escalated enforcement, with dispersion parameter between 0 and 0.1.

[27]For comparison, I have also estimated the non-normalized regressions dropping these observations. Results, shown in Appendix C, are similar to the main results in panel A of table 3.

the continuous variables, the left-hand side variable is divided by capacity factor.

Results for both normalized and non-normalized outcome variables are given in table 3. Panel A shows the total effect of divestiture on safety, whereas panel B shows the effect for a given capacity factor. To compare the magnitude in the OLS specifications with the magnitude in the count specifications, I have shown the percentage change in the expected number of counts attributable to divestiture for all regressions.[28] For all five of the safety measures, the coefficient on divestiture is negative in panel A. For initiating events, the coefficient is -0.19; for fires, the coefficient is -0.62; and for escalated enforcement, the coefficient is -0.43. For collective worker radiation exposure, divestiture is associated with a drop of 42 person-rems; average exposure drops by 0.03 rems. While the point estimates are not precisely estimated, the magnitude of the coefficient is economically significant for all five measures. For initiating events, for instance, divestiture leads to a 17 percent reduction in the expected monthly event count. For fires, the change is -46 percent, and for escalated enforcement the change is -35 percent. Furthermore, some moderate positive effects can be ruled out at the 5 percent level: for initiating events and escalated enforcement, the upper bound of the 95 percent confidence interval is 0.06.

When the dependent variable is scaled by capacity factor (panel B), the coefficient on divestiture is more negative. Divestiture leads to a 28 percent change in initiating events for a given capacity factor, and the coefficient is statistically significant at the 1% level. For fires, the change in expected value is 54 percent (significant at the 10 percent level), and for escalated enforcement, 42 percent (significant at the 5 percent level). For the two worker radiation exposure variables, the effect is even larger, but it is not precisely estimated.

Overall, it appears that divestitures did not lead to worsened safety records, and they may have led to some decreases in unsafe events. Moreover, divestitures increased generation, and controlling for this, safety substantially improved. Both the total effect on safety (when unscaled by capacity factor) and the effect controlling for generation are of policy interest. As such, tables throughout this paper provide estimates for both outcomes.

These results match anecdotal evidence that deregulation led to improved safety. Whereas the NRC had expressed concerns about plant safety following deregulation, a regional administrator said in 2001 that "[m]ost people have gotten the understanding if you do it right the first time, and you emphasize safety and managing things better, it has a positive effect on the bottom line."[29]

1.4.3 Simultaneity between Safety and Generation

Ideally, one would treat the simultaneity between safety and generation as a full system of equations to estimate the direct effect of divestiture on each. To understand this simultaneity, consider two cases. First, if a fire occurs in the turbine area, the plant must shut down until repairs can be made; in this case, unsafe events lead to lower generation. On the other

[28] For the count specifications, the percentage change in the expected number of counts is equal to $\exp(\beta) - 1$. For the OLS specifications, the percentage change in the expected number of counts is equal to $\frac{\beta}{E[y_{it}|d_{it}=0;\alpha_i,v_t]}$.

[29] Source: Wald, Matthew L. 2001. "Despite Fear, Deregulation Leaves Nuclear Reactors Working Harder, Longer and Safer." *New York Times*, February 18.

hand, if a plant shuts down for some exogenous reason, it is less likely to have a fire, because the turbine is not moving. In this case, increased generation leads to more unsafe events.[30] Throughout this section, I focus on initiating events and fires, for which this intuition is most applicable.

Unfortunately, because I do not observe the generation level at a plant prior to a fire, the direct effect of generation on safety cannot be observed separately from the direct effect of safety on generation. While I do observe the total generation for a month, this is conditional on whether a fire occurred.

The full system of equations is

$$\begin{aligned} s &= f(d, g, X) \\ g &= k(d, s, X) \end{aligned}$$

Here s is an unsafe event, g is generation, d is a divestiture dummy (the variable of interest), and X is a vector of exogenous variables. The direct effect of divestiture on each endogenous variable cannot be estimated econometrically for this system, unless there is an instrumental variable for each equation. Unfortunately, there are no credible candidates for such instruments. Refueling outages, for instance, might affect unsafe events only through their impact on generation, but refueling outages occur at the same time as other planned maintenance, which is certainly correlated with safety.

1.4.3.1 Calculating the Direct Effect of Divestiture on Safety

However, by making certain assumptions, the direct effect of divestiture on safety can be calculated from this system. Intuitively, the direct effect of divestiture on unsafe events could be positive or negative, but divestiture also increases generation, thereby increasing the exposure of the plant to an event. Then the direct effect on divestiture will be more negative, or less positive, than the total effect. Consider the total derivative $\frac{df}{dd} \cdot \frac{1}{s} = \frac{\partial f}{\partial d} \cdot \frac{1}{s} + \left(\frac{\partial f}{\partial g} \cdot \frac{g}{s}\right) \cdot \left(\frac{dg}{dd} \cdot \frac{1}{g}\right)$.[31] We want to know $\frac{\partial f}{\partial d}$, the direct effect, whereas what was estimated previously was $\frac{df}{dd}$, the total effect. Taking the preferred empirical estimate from Davis and Wolfram (2012), assume that $\frac{dg}{dd} \cdot \frac{1}{g} = 0.10$; divestiture increases generation by approximately 10 percent.[32] Also, make the neutral assumption that $\frac{\partial f}{\partial g} \cdot \frac{g}{s} = 1$; a one percent increase in

[30]Note that this analysis, which focuses on the difference in exposure when a plant is on versus off, does not account for the difference in exposure during plant ramp up and ramp down. If divested plants increase their generation time but decrease their ramping times, their exposure to a safety event could, on net, fall.

[31]For notational simplicity, I drop the year and fixed effects, which are the only exogenous variables other than divestiture. For this simplification to be valid, I assume that divestiture does not impact either the time-invariant reactor effects or the reactor-invariant year effects.

[32]Note that the Davis and Wolfram estimate is also for the total effect, which includes the indirect effect divestiture on generation through safety. However, the difference between the direct and total effects in this case are likely small, since unsafe events are infrequent. Accordingly, assume a direct effect of 10 percent for now; the difference between the direct and indirect effects are explored below.

generation time leads to an expected increase in unsafe events of one percent.[33] Finally, recall that the total effect of divestiture $\frac{\partial f}{\partial d} \cdot \frac{1}{s}$ is empirically estimated to be a reduction of 17 percent for initiating events and 46 percent for fires.[34] Then the direct effect of divestiture on unsafe events is calculated to be -0.27 for initiating events and -0.56 for fires.[35] Thus while divestiture leads to a total effect of a reduction of 17 percent in initiating events, the direct effect is a reduction of 27 percent. The difference arises from the indirect effect through generation. These results do not change much when the relationship between generation and unsafe events is allowed to vary. For $\frac{\partial f}{\partial g} \cdot \frac{g}{s} = 0.5$, the direct effect of divestiture is -0.22 for initiating events and -0.51 for fires; for $\frac{\partial f}{\partial g} \cdot \frac{g}{s} = 1.5$, the direct effects are -0.32 and -0.61. Note that these estimates are very similar to the normalized estimates calculated in the previous section (-28 percent and -54 percent).[36]

1.4.3.2 Calculating the Direct Effect of Divestiture on Generation

A similar exercise can be performed for the effect of divestiture on generation. As described above, this is likely to be very close to the total effect: there are few unsafe events in any given month, so the indirect effect of these incidents on generation is likely to be small. Suppose the elasticity of generation with respect to initiating events is -0.016: a one percent increase in events leads to an expected decrease in generation of 0.016 percent. This assumed elasticity is derived from (1) noting that initiating events only occur in approximately 10 percent of months, and (2) assuming that an incident leads to five days of lost generation time, i.e., 13 percent of the month's generation. Similarly, the elasticity of generation with respect to fires is -0.002, from noting that fires occur in 0.7 percent of months and assuming eight days of lost generation time.[37] Then for total effect of divestiture on generation of 10 percent, the direct effect after accounting for both fires and initiating events is calculated to be 9 percent.[38] This is very close to the total effect of 10 percent, because unsafe events occur fairly infrequently.

[33]The elasticity could be smaller if increased generation time allows for built-up expertise. On the other hand, the elasticity could be larger if there is fatigue, for instance, of employees as generation time increases.

[34]The relevant statistics from table 3 are not the raw coefficients from each regression, but rather the percentage change in expected value.

[35]Block-bootstrapped standard errors (clustered at the plant level) that account for the correlation between effect on generation and the effect on safety are 0.09 for initiating events and 0.30 for fires.

[36]Robustness checks to account for differential historical usage give similar results. For instance, I can include cumulative lifetime generation as an exogenous variable (excluding current generation, which is subject to simultaneity); this conditions on relative plant usage. As shown in Appendix C, results for the divestiture coefficient are nearly identical. Rather than using cumulative generation, several lags of generation can be included as right-hand side variables. As shown in Appendix C, this gives similar results. The coefficient on divestiture is somewhat smaller in absolute value, but is not statistically different from the main results.

[37]I examined daily generation data and descriptions (from Davis and Wolfram 2012) for twenty randomly selected fires and twenty randomly selected initiating events. The mean number of days with generation below 50 percent of capacity following the event was four for initiating events and seven for fires. There were typically a few more days of ramping with generation levels slightly lower than 100 percent of capacity.

[38]The block-bootstrapped standard error (clustered at the plant level) that accounts for the correlation between effect on generation and the effect on safety is 0.02.

1.4.4 Robustness Checks

Several robustness checks give very similar results (table 4). First, I estimate a conditional negative binomial with fixed effects specification in columns (1), (4), and (7).[39] For this specification, the individual effects α_i enter the conditional negative binomial specification only in the variance parameter. That is, this specification does not allow for heterogeneity in the mean across units (Allison and Waterman 2002). Next, columns (2), (5), and (8) give results for a Poisson specification. All point estimates and standard errors are very similar to the results given by the unconditional negative binomial model. Finally, I show OLS specifications in columns (3), (6), and (9). To compare the magnitude in the OLS specifications with the magnitude in the count specifications, I show the percentage change in the expected number of counts attributable to divestiture for all regressions. Overall, the results are stable to various assumptions on functional form. For all future regressions using the three count variables, I show results for the unconditional negative binomial specification.

I also examine whether the results are driven by outliers. I perform a jackknife procedure both at the plant level and the year level. As shown in Appendix C, the results are stable to dropping any one plant or any one year. Additionally, I show that the results are not driven by the company (Exelon) with the greatest number (17) of divested reactors. Results in Appendix C when Exelon reactors are dropped are very similar to the main results, with the exception of the worker radiation exposure measures.

1.4.5 Heterogeneity

I next explore whether heterogeneity can also be observed across reactor fixed characteristics (table 5). I first divide reactors according to type (BWR versus PWR). Since BWR reactors were more likely to divested, one might worry about either bias from selection or about external validity. With the exception of the worker radiation exposure equations (columns 4 and 5), the coefficient on divestiture is not statistically different for BWR versus PWR reactors. I next divide reactors by age, defining newer reactors (51 of 103) as those entering commercial operations in 1979 or later. Finally, I divide by design capacity, defining large reactors (49 of 103) as those with current capacity of at least 1000 MW. Age and size are not correlated with divestiture (table 2), but heterogeneity in the effect of divestiture does appear. There is some evidence that newer and larger reactors improved more, particularly for initiating events and escalated enforcement. The size and age definitions are highly correlated: the majority of newer reactors are large, and the majority of older reactors are small. As such, there is unfortunately insufficient power to separately test the effects of size and age. All heterogeneity results (in Appendix C) are similar when the dependent variable is normalized by capacity factor.

[39] The fixed-effects conditional negative binomial model begins with a Poisson specification and then assumes the Poisson parameter follows a $gamma(\exp(x_{it}B), \alpha_i)$ distribution. This implies that the variance is proportional to the mean. The α_i parameter is allowed to vary by reactor in the fixed-effects specification.

1.4.6 State-Level Selection

Next I exclude a series of states to address potential selection concerns. First I exclude Michigan, where some but not all reactors were divested; in all other states, either all or none were divested. Second, I exclude California, where fossil fuel plants but not nuclear plants were divested. Furthermore, one of the nuclear plants (Diablo Canyon) is subject to strong incentive regulations. Third, I exclude Iowa, Vermont, and Wisconsin, where reactors were divested but the electricity market was not deregulated. Finally, I exclude the Northeast, where most divestitures occurred, to see if unobserved regional differences drive the results. For all four specifications (table 6), the results are robust. The coefficient on divestiture is almost always negative, and the magnitudes are largely unchanged from the main specification. One exception is collective worker radiation exposure, which is sensitive to excluding the Northeast.[40]

1.4.7 Spillovers and Consolidation

Previous work has shown spillovers of safety practices across plants, including to the companies operating non-divested plants (MIT 2003, Rees 1994). There are several organizations that facilitate knowledge-sharing across the plants: the World Association of Nuclear Operators (WANO), the U.S.-based Institute of Nuclear Power Operations (INPO), the Electric Power Research Institute (EPRI), and EUCG. As described in section 1.2.2, INPO in particular has had substantial impact on the industry by facilitating best-practices sharing across all reactors in the U.S. (Rees 1994). If the owners of divested plants share their practices with the owners of non-divested plants, the regression results above will give a lower bound on the overall effect of divestiture. The control group (non-divested plants) will have been impacted by divestiture, implying a poor counterfactual. If the control group improves following divestiture, the coefficient on the divestiture dummy will be smaller than the true effect on the divested plants. Additionally, the regression results will fail to capture the effect on the non-divested plants. It is not possible to test for these spillovers across all plants using this paper's empirical strategy. There is some suggestive evidence that this has occurred; for instance, safety records have improved nationwide in the last decade. This could also, however, be the result of other changes, such as more stringent NRC regulations. In Appendix C, results are given for a test of intra-firm spillovers between divested and non-divested plants. I generally do not find an effect, with the exception of escalated enforcement, which falls at non-divested plants owned by companies that also own plants in regulated environments. One possible explanation is that operators fear scrutiny at all of

[40] As described in section 1.2.1, another selection concern relates to the timing of divestiture. If, for instance, plants that expected to have larger gains following divestiture were sold first, my results would be weighted in favor of those plants. In Appendix C, I examine the robustness of the main results to including only four years of post-divestiture data at each of the plants. Results are noisy but similar to results including all years of data. Additionally, I examine the robustness of the results to using deregulation dates rather than divestiture dates. It is not clear what date to use, and related papers have used several measures of deregulation dates (Fabrizio, Rose, and Wolfram 2007, Craig and Savage 2009). Appendix C shows results for four different measures of deregulation, all of which will introduce measurement error. Results are generally robust to these alternative dates, with the exception of initiating events.

their plants if an escalated enforcement action is taken at one plant. Although no evidence of intra-firm spillovers is found, it is possible there are nation-wide spillovers, biasing the main results. Finally, I also test for a consolidation effect, measured as the number of reactors owned by a plant's parent company, but generally fail to find an effect.

1.4.8 Dynamic Effects

There is some evidence that the benefits of divestiture would increase over time, both because some plant modifications would take time, and because the companies would learn. The event study graph (figure 1) showed a change in the trend of safety records at divested relative to non-divested plants. Accordingly, I add linear trends pre- and post-divestiture at divested plants, with results in table 7.[41] This specification also allows me to look for differential trends prior to divestiture. The coefficients on the linear trends are scaled to represent a three-year change. The pre-divestiture trend is generally very small, indicating that the plants were on similar trajectories.[42] Overall, trends post-divestiture are negative, consistent with learning. These downward trends are reassuring that the results are not driven by temporary changes following divestiture.

1.5 Conclusion

This paper provides empirical evidence on the effects of divestiture on nuclear power plant safety in the United States. I examine both the total effect of divestiture on safety and the effect when controlling for increased generation levels. The total effect is composed of both the direct effect on safety and an indirect effect. The latter arises from the fact that generation increased following divestiture, and thus plants may have experienced an increase in exposure to unsafe events. The total drop in safety incidents is estimated to be 17 percent for initiating events, 46 percent for fires, and 35 percent for escalated enforcement. While none of these effects is statistically significant at the 5 percent level, moderate positive effects can be ruled out at the 5 percent level. Worker radiation exposure, measured either collectively or on average, also decreases. When controlling for generation increases, I find that the direct effect on safety is more negative and more statistically significant. For a given capacity factor, the drop in safety incidents is estimated to be 28 percent for initiating events, 54 percent for fires, and 42 percent for escalated enforcement. Results are also larger, although imprecisely estimated, for worker radiation exposure.

The results are similar for a number of robustness checks, including concerns about selections on technology and location, the inclusion of pre-treatment trends, and jackknife procedures. In extensions, I find some heterogeneity, with larger results for newer, bigger reactors. I do not find evidence of spillovers or consolidation. However, it is likely that spillovers in the form of best-practices sharing exist, implying that deregulation had a larger

[41] Ideally, one would estimate a full event study for all empirical specifications. However, for these infrequent events the results are extremely noisy. The coefficients are given in Appendix C.

[42] One exception is escalated enforcement, which shows an upward trend prior to divestiture: the plants that were eventually sold off were worsening their safety records.

effect than the results given above. Finally, a specification allowing for differential trends indicates that the effect has grown larger over time, providing reassurance that the results are not driven by temporary changes.

Several caveats apply. First, as described in section 1.3.1, the available information on power plant safety does not directly measure the risk of catastrophic failure. The measures analyzed are, however, widely used as indicators of power plant safety. A related concern about the data used is that the incentive alignment described applies only when maintenance for safety is positively correlated with a plant's ability to generate. This concern is mitigated as long as the NRC can force a plant to close. A third concern, then, is the possibility of either incomplete NRC enforcement or regulatory capture. Also, there is the possibility of end-of-life problems when plants close: if the plant operator can no longer earn a future stream of operating profits, it may choose to forgo safety-related repairs. Finally, it should be noted that the estimates in this paper are for the effect of the treatment on the treated units. That is, if plants that were never deregulated are dissimilar in time-invariant ways from the regulated plants, it is possible they would respond to divestiture differently.

Overall, this paper speaks to a number of timely issues, including the changing structure of the electricity industry and the incentives for safety at nuclear power plants. Although intuition is given throughout this paper for some of the mechanisms at work, theoretical predictions of the effect of deregulation on plant safety are not possible. As such, an empirical analysis of the effect is the best evidence available. While the infrequency of unsafe events at nuclear plants makes precise statistical estimates difficult, the results match anecdotal evidence. Deregulation of electricity markets led to increased operating efficiency, and it did not come at the cost of plant safety.

1.6 Tables for Chapter 1

Table 1: Annual Reactor-Level Summary Statistics

	Mean	Std. Dev.	Min	Max
A. Safety measures:				
Initiating events	0.86	1.07	0	6
Fires	0.07	0.27	0	2
Collective worker radiation exposure (person-rems)	116.04	89.95	1.40	893.01
Average worker radiation exposure (rems)	0.15	0.06	0.01	0.47
Escalated enforcement	0.44	0.75	0	6
B. Reliability measures:				
Generation (million MWh)	7.27	2.13	-0.12	11.77
Capacity factor	0.88	0.16	-0.01	1.20

Notes: Data are for 103 nuclear power reactors operating in the U.S. from 1996-2009. Both radiation exposure variables are measured at the plant level. For collective exposure, the numbers in this table are a simple mean across units within the plant. Some plants (e.g., Browns Ferry) are dropped from the sample because radiation exposure measurements include closed units. Also, data on these variables is only available through 2008. Capacity factor is defined as generation divided by design capacity. Generation is net, not gross, and accordingly can take on negative values. Capacity factor can similarly be negative. It can also be greater than 1 because of changes to reactor capacity over time (uprates). N = 1442 for count variables and reliability measures, 1259 for radiation variables.

Table 2: Comparing Divested and Non-Divested Nuclear Reactors

	never divested	later divested	t-stat	p-value
A. Safety measures:				
Initiating events	1.05	0.92	0.93	0.36
Fires	0.085	0.056	0.82	0.41
Collective worker radiation exposure (person-rems)	148.38	163.73	-0.84	0.41
Average worker radiation exposure (rems)	0.18	0.19	-0.63	0.53
Escalated enforcement	0.75	0.98	-1.54	0.13
B. Reliability measures:				
Net generation (million MWh)	6.90	5.67	2.17	0.03
Capacity factor	0.82	0.70	2.59	0.01

Notes: Data are for the 103 nuclear power reactors operating in the U.S. from 1996-1998, by eventual divestiture status: independent power producers versus regulated investor-owned utilities. For collective exposure, the numbers in this table are a simple mean across units within the plant. Some plants (e.g., Browns Ferry) are dropped from the sample because radiation exposure measurements include closed units. For the count variables and reliability measures, (measured by reactor), N = 165 for never divested units, 144 for later divested units. For the radiation exposure variables (measured plant), N = 95 for never divested plants, 87 for later divested plants. One reactor (Watts Bar 1) starts commercial operation during this time. T-tests are clustered at the plant level.

Table 3: The Effect of Divestiture on Nuclear Power Plant Safety

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforcement	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
Divestiture	-0.192	-0.622	-0.426*	-42.2	-0.025
	(0.130)	(0.433)	(0.247)	(67.3)	(0.022)
Change in expected value	-17%	-46%	-35%	-25%	-18%
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2245	1950	1442	1749	1749
B: Dependent variable is normalized by capacity factor					
Divestiture	-0.335***	-0.767*	-0.552**	-180.2	-0.108
	(0.122)	(0.440)	(0.277)	(278.3)	(0.103)
Change in expected value	-28%	-54%	-42%	-93%	-69%
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2207	1925	1425	1729	1729

Notes: Observation is a commercial nuclear power reactor (U.S.) in a year for the left-most three columns and a commercial nuclear power plant in a year for the right-most two columns. Divestiture is a dummy variable equal to 1 if the reactor is owned by an independent power producer, and 0 if the reactor is owned by a regulated investor-owned utility. Normalization for the count regressions is accomplished by including capacity factor as an independent variable with coefficient constrained to unity. In columns (1), (2), and (3), the percentage change in expected value is equal to exp(coefficient) minus one; for columns (4) and (5), it is equal to the coefficient divided by the sample average at non-divested reactors. Samples dates vary by variable. Initiating events are 1988-2009; fires are 1991-2009; escalated enforcement is 1996-2009; and radiation exposure is 1974-2008. For fires and escalated enforcement, some reactors (34 and 2, respectively) are dropped because all observations are zero. Additionally, some observations have zero capacity factor and are dropped in panel B. Standard errors are clustered by plant. Stars (*, **, and ***) denote 10%, 5%, and 1% significance.

Table 4: Robustness Checks: The Effect of Divestiture on Nuclear Power Plant Safety

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	Initiating Events			Fires			Escalated Enforcement		
A: Dependent variable is not normalized									
Divestiture	-0.21	-0.18	-0.19	-0.60	-0.62	-0.04	-0.41	-0.43*	-0.22*
	(0.13)	(0.13)	(0.13)	(0.45)	(0.43)	(0.03)	(0.26)	(0.25)	(0.12)
Change in expected value	-19%	-16%	-15%	-45%	-46%	-48%	-34%	-35%	-56%
Specification	CNB	Poiss	OLS	CNB	Poiss	OLS	CNB	Poiss	OLS
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2245	2245	2245	1950	1950	1950	1442	1442	1442
B: Dependent variable is normalized by capacity factor									
Divestiture	-0.34**	-0.32**	-0.54**	-0.74	-0.77*	-0.07	-0.53*	-0.55**	-0.37
	(0.12)	(0.12)	(0.23)	(0.46)	(0.44)	(0.05)	(0.28)	(0.28)	(0.31)
Change in expected value	-29%	-27%	-31%	-52%	-54%	-68%	-41%	-42%	-67%
Specification	CNB	Poiss	OLS	CNB	Poiss	OLS	CNB	Poiss	OLS
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2207	2207	2207	1925	1925	1925	1425	1425	1425

Notes: Observation is a commercial nuclear power plant (U.S.) in a year. Divestiture is a dummy variable equal to 1 if the reactor is owned by an independent power producer, and 0 if the reactor is owned by a regulated investor-owned utility. CNB is a conditional negative binomial; OLS is ordinary least squares. Results are nearly identical with an unconditional negative binomial with constant dispersion parameterization. Normalization for the count regressions is accomplished by including capacity factor as an independent variable with coefficient constrained to unity. For the count specifications, the percentage change in expected value is equal to exp(coefficient) minus one; for OLS, it is equal to the coefficient divided by the mean number of counts at non-divested reactors. Sample dates vary by variable. Initiating events are 1988-2009; fires are 1991-2009; and escalated enforcement is 1996-2009. For fires and escalated enforcement, some reactors (34 and 2, respectively) are dropped in the count regressions because all observations are zero. Additionally, some observations have zero capacity factor and are dropped in panel B. Standard errors are clustered by plant. Stars (*, **, and ***) denote 10%, 5%, and 1% significance.

Table 5: Heterogeneity by Reactor Characteristics

Dependent variable:	(1) Initiating Events	(2) Fires	(3) Escalated Enforcement	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
	Dependent variable is not normalized				
Divestiture, BWR	-0.13	-0.57	-0.44	-166.5*	-0.065**
	(0.13)	(0.52)	(0.30)	(87.5)	(0.029)
Divestiture, PWR	-0.26	-0.70	-0.41	109.9**	0.024
	(0.20)	(0.62)	(0.31)	(48.8)	(0.022)
Chi-squared stat	0.41	0.03	0.01	12.13***	6.81**
Divestiture, older reactors	-0.01	-0.29	-0.22	-78.3	-0.045
	(0.15)	(0.44)	(0.25)	(89.7)	(0.028)
Divestiture, newer reactors	-0.40**	-1.14	-0.76**	14.5	0.006
	(0.18)	(0.68)	(0.33)	(47.7)	(0.021)
Chi-squared stat	4.01**	1.50	3.44*	1.42	2.95*
Divestiture, small reactors	0.02	-0.61	-0.27	-54.2	-0.060*
	(0.15)	(0.51)	(0.24)	(91.3)	(0.030)
Divestiture, large reactors	-0.41**	-0.63	-0.64*	-29.4	0.012
	(0.16)	(0.61)	(0.35)	(75.7)	(0.021)
Chi-squared stat	5.16**	<0.01	1.52	0.06	4.82*
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2245	1950	1442	1749	1749

Notes: A separate regression is run for each heterogeneous effect (PWR versus BWR, reactor vintage, and reactor size). Observation is a commercial nuclear power reactor (U.S.) in a year for the left-most three columns and a commercial nuclear power plant in a year for the right-most two columns. Divestiture is a dummy variable equal to 1 if the reactor is owned by an independent power producer, and 0 if the reactor is owned by an investor-owned utility. I define newer reactors (51 of 103) as those entering commercial operations in 1979 or later. I define large reactors (49 of 103) as those with current capacity of at least 1000 MW. Initiating events, fires, and escalated enforcement are count variables. Collective worker radiation exposure is measured in person-rems, and average worker radiation exposure in rems. Samples dates vary by variable. Initiating events are 1988-2009; fires are 1991-2009; escalated enforcement is 1996-2009; and radiation exposure is 1974-2008. For fires and escalated enforcement, some reactors (34 and 2, respectively) are dropped in the count regressions because of all zero outcomes. Standard errors are clustered by plant. Stars (*, **, and ***) denote 10%, 5%, and 1% significance.

Table 6: State-Level Selection

Dependent variable: Divestiture	(1) Initiating Events	(2) Fires	(3) Escalated Enforcement	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
A. Dependent variable is not normalized					
Excluding Michigan	-0.16	-0.63	-0.42*	-32.9	-0.020
	(0.13)	(0.44)	(0.25)	(69.3)	(0.022)
Excluding California	-0.20	-0.63	-0.44*	-43.0	-0.025
	(0.13)	(0.44)	(0.25)	(68.2)	(0.022)
Excluding Iowa, Vermont, and	-0.21	-0.73*	-0.48*	-50.0	-0.030
Wisconsin	(0.13)	(0.44)	(0.26)	(70.8)	(0.023)
Excluding Northeast	-0.33*	-1.05*	-0.33	16.4	-0.017
	(0.19)	(0.63)	(0.29)	(73.8)	(0.040)
B: Dependent variable is normalized by capacity factor					
Excluding Michigan	-0.31**	-0.77*	-0.56*	-192.4	-0.143
	(0.12)	(0.45)	(0.29)	(292.9)	(0.108)
Excluding California	-0.34***	-0.77*	-0.56**	-178.5	-0.106
	(0.12)	(0.45)	(0.28)	(281.3)	(0.105)
Excluding Iowa, Vermont, and	-0.36***	-0.89**	-0.63**	-210.3	-0.126
Wisconsin	(0.12)	(0.45)	(0.30)	(294.6)	(0.108)
Excluding Northeast	-0.48***	-1.28*	-0.47	44.7	-0.082
	(0.18)	(0.68)	(0.31)	(165.1)	(0.069)

Notes: Each coefficient is from a separate regression (eight per outcome variable). Observation is a commercial nuclear power reactor (U.S.) in a year for the left-most three columns and a commercial nuclear power plant in a year for the right-most two columns. Divestiture is a dummy variable equal to 1 if the reactor is owned by an independent power producer, and 0 if the reactor is owned by an investor-owned utility. Columns (1), (2), and (3) are negative binomial specifications with year and reactor effects. Columns (4) and (5) are OLS specifications with year and facility effects. Samples dates vary by variable. Initiating events are 1988-2009; fires are 1991-2009; escalated enforcement is 1996-2009; and radiation exposure is 1974-2008. For fires and escalated enforcement, some reactors (34 and 2, respectively) are dropped in the count regressions because of all zero outcomes. Additionally, some observations have zero capacity factor and are dropped in panel B. Standard errors are clustered by plant. Stars (*, **, and ***) denote 10%, 5%, and 1% significance.

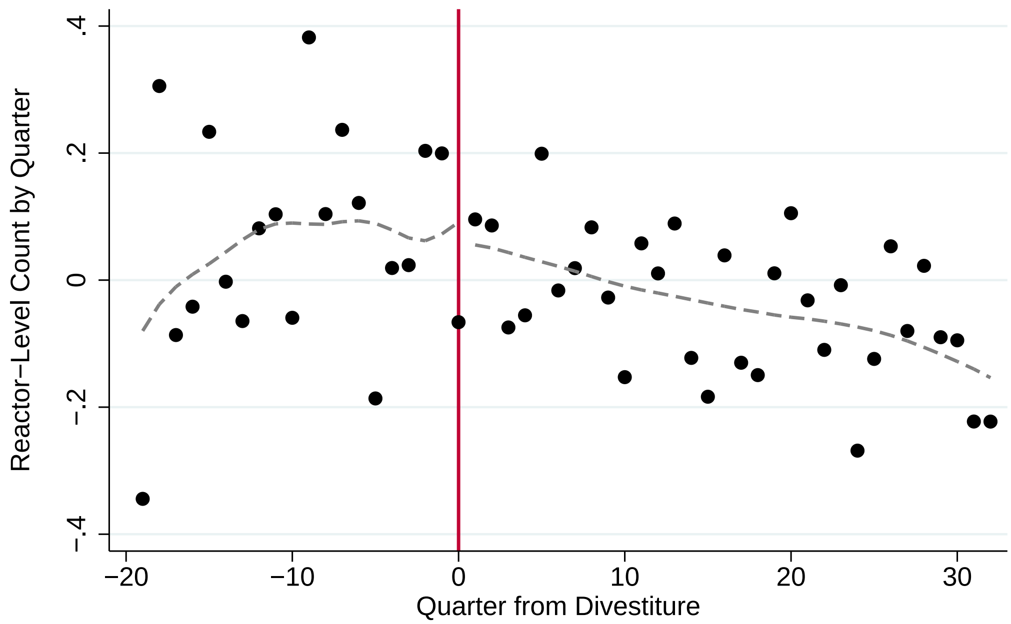
Table 7: Learning

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforcement	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
Divestiture	-0.01	-0.01	-0.43	-38.7	0.030
	(0.18)	(0.68)	(0.32)	(42.2)	(0.025)
Linear trend pre-divestiture	-0.01	-0.06	0.35**	16.4	-0.013
	(0.04)	(0.16)	(0.15)	(24.5)	(0.011)
Linear trend post-divestiture	-0.14	-0.38	-0.29*	-46.5*	-0.016
	(0.12)	(0.35)	(0.17)	(26.3)	(0.010)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2245	1950	1442	1749	1749
B. Dependent variable is normalized by capacity factor					
Divestiture	-0.14	-0.07	-0.41	20.5	0.277
	(0.17)	(0.68)	(0.35)	(227.7)	(0.254)
Linear trend pre-divestiture	0.01	-0.07	0.27*	-26.6	-0.119
	(0.04)	(0.16)	(0.16)	(100.1)	(0.088)
Linear trend post-divestiture	-0.19	-0.42	-0.35**	-108.8	-0.025
	(0.12)	(0.35)	(0.17)	(106.3)	(0.050)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2207	1925	1425	1729	1729

Notes: Learning variable has been scaled to represent a three-year change. Samples dates vary by variable. For fires and escalated enforcement, some reactors (34 and 2, respectively) are dropped in the count regressions because of all zero outcomes. Additionally, some observations have zero capacity factor and are dropped in panel B. Standard errors are clustered by plant. Stars (*, **, and ***) denote 10%, 5%, and 1% significance.

1.7 Figures for Chapter 1

Figure 1: Effect of Divestiture on Unsafe Events, Quarterly Event Study



Notes: This figure plots unsafe events at divested units relative to non−divested units. Time is normalized relative to divestiture. Unsafe events include initiating events, fires, and escalated enforcement. The median divestiture is in 2001. Quarter−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Appendix A: Model

I model profit maximization of a nuclear power plant in a competitive generation market, then I derive implications for expenditures on reliability and safety maintenance.

Profit-Maximization with Reliability

Consider a baseload nuclear power plant in a deregulated electricity generation market. For simplification, assume the power plant has only one reactor. The power plant faces a given price per megawatt-hour (MWh) p and given fuel and other variable costs per MWh, c^o. The market price of electricity generation is determined by the marginal cost of the marginal plant. Variable costs for nuclear plants are lower than for fossil fuel plants, implying that nuclear plants are not the marginal plants. According to a recent EIA report (EIA 2009), variable costs are 2.17 cents per kilowatt-hour for nuclear plants and 4.05 for fossil-steam plants. First, assume that the nuclear plant is a price-taker.[43] Second, assume that $p > c^o$; the market price is higher than the nuclear plant's variable costs.[44]

If the plant is operating, it operates at capacity, i.e., producing quantity q of electricity. Let operating (not total) profits $\pi = pq - c^o q$. Assume there are no ramping or start-up costs. The plant can choose some level of maintenance a to purchase; thus a is an endogenous effort variable. Increases in a can be thought of as increases in either the quantity or quality of effort. Most maintenance for nuclear power plants requires the plant to be offline, so maintenance incurs both direct costs and lost operating profits. The cost of maintenance is $c(a, \pi)$, where $c(a, \pi) \geq 0$, $\frac{\partial c}{\partial a} > 0$, $\frac{\partial c}{\partial \pi} > 0$, $\frac{\partial^2 c}{\partial^2 a} > 0$ and $\frac{\partial^2 c}{\partial a \partial \pi} > 0$. The intuition for the assumptions on the first and second derivatives with respect to operating profits π is that additional maintenance requires a longer time offline, so more revenue is lost.[45] In any given period, there is a probability $r(a) \in (0, 1)$ that the plant will experience an unplanned outage (or "scram" or "trip"), conditional on the plant deciding ex-ante to operate. Then the probability of being able to operate as planned is given by $1 - r(a)$. Assume $r'(a) < 0$: maintenance (effort) decreases the probability of an unplanned outage. Also, $r''(a) > 0$: the probability decreases at a decreasing rate. Intuitively, the probability asymptotes as maintenance increases. In the event of an unplanned outage, the firm earns no revenue (as it produces no electricity) and incurs additional costs $c^u > 0$. These additional costs may include repair work, increased (safety) regulatory scrutiny, or bad publicity. The firm's profit

[43] There is potential for the owner of a nuclear power plant to exercise market power, if it owns other generators. However, if the other generators have higher marginal costs than the nuclear plant, exercising market power by shutting down the nuclear plant is not the first-best strategy of the firm. Rather, the firm would take the higher cost plant offline. Moreover, if the nuclear power plant is baseload, the owner may be required to purchase replacement power when the plant is down. Since the replacement power is more costly than the nuclear plant's generation, the firm has no incentive to exercise market power by taking the nuclear plant offline.

[44] For representative supply and demand curves showing nuclear marginal costs compared to fossil fuel costs, see Griffin and Puller (2005).

[45] It is straightforward to consider the case where $\frac{\partial c}{\partial \pi} = 0$, i.e., maintenance does not require the plant to be offline.

maximization problem is[46]

$$\max_{a} \quad (1-r(a))\cdot\pi - r(a)\cdot c^u - c(a,\pi) \tag{1.1}$$

The first-order condition is

$$-r'(a)\cdot\pi - r'(a)\cdot c^u - \frac{\partial c(a,\pi)}{\partial a} = 0 \tag{1.2}$$

The firm chooses the level of maintenance a such that the marginal benefit of an additional unit of maintenance $-r'(a)\cdot\pi - r'(a)\cdot c^u$ equals the marginal cost $\frac{\partial c(a,\pi)}{\partial a}$. The marginal benefit of an additional unit of maintenance is an increased likelihood of earning revenue and a decreased likelihood of paying for an unplanned outage. Comparative statics on the exogenous revenue and cost variables is straightforward. By the implicit function theorem,

$$\begin{bmatrix} \frac{\partial a}{\partial \pi} \\ \frac{\partial a}{\partial c^u} \end{bmatrix} = -\left[-r''(a)\cdot\pi - r''(a)\cdot c^u - \frac{\partial^2 c(a,\pi)}{\partial^2 a}\right]^{-1}\cdot\begin{bmatrix} -r'(a) - \frac{\partial^2 c}{\partial a\partial\pi} \\ -r'(a) \end{bmatrix} \tag{1.3}$$

At the profit maximizing level of a, $\left[-r''(a)\cdot\pi - r''(a)\cdot c^u - \frac{\partial^2 c(a,\pi)}{\partial^2 a}\right]$ is negative (by the second order condition, which is satisfied according to the above assumptions),[47] and recall that $r'(a)$ is assumed to be negative and $\frac{\partial^2 c(a,\pi)}{\partial^2 a}$ positive. The sign on $\frac{\partial a}{\partial \pi}$ is indeterminate; both planned maintenance outages and unplanned outages lead the firm to lose revenue. If one instead assumes that maintenance does not require the plant to be offline, i.e., $\frac{\partial c}{\partial \pi} = 0$, then maintenance a is increasing in potential revenue. (Note that all results on $\frac{\partial a}{\partial \pi}$ imply the same result on $\frac{\partial a}{\partial p}$, since $\frac{\partial \pi}{\partial p} = 1$.) The sign on $\frac{\partial a}{\partial c^u}$ is positive; maintenance is increasing in the cost of an unplanned outage.

Profit-Maximization with Reliability and Safety

The above model considers plant reliability rather than safety. Suppose that the probability of an unsafe event is $s(a) \in (0,1)$ with $s'(a) < 0$ and $s''(a) > 0$; that is, the same maintenance actions that improve reliability also improve safety. Suppose the total cost of

[46] As an alternative way to see how maintenance costs depend on operating profits, re-write the firm's total profits as $(1-r(a)-p(a))\cdot\pi - r(a)\cdot c^u - c(a)$, where $r(a) \in (0,1)$ is the probability of an unplanned outage, and $p(a) \in (0,1)$ is the fraction of time spent on planned outages. Thus all time is spent on either planned outages, unplanned outages, or generation. As before, $r'(a) < 0$, $c'(a) > 0$, and now $p'(a) > 0$: the time spent on a planned outage is increasing in the amount of maintenance done. Rearranging the firm's total profit function gives $(1-r(a))\cdot\pi - r(a)\cdot c^u - p(a)\cdot\pi - c(a)$. Let $\tilde{c}(a,\pi) = p(a)\cdot\pi + c(a)$, so that profits equal $(1-r(a))\cdot\pi - r(a)\cdot c^u - \tilde{c}(a,\pi)$. The latter expression is the same as equation 1.1, showing how the cost of maintenance depends on operating profits.

[47] The key assumption for satisfying the second order condition is that $r''(a) > 0$. Intuitively, this is satisfied for large a if the probability of an unplanned outage asymptotes towards zero as maintenance increases. If $r(a)$ is S-shaped, with $r''(a) < 0$ for small values of a, there could be a corner solution with no maintenance. All that is necessary to rule out this case is to assume that the optimal a is beyond the inflection point; alternatively, one could assume that the regulatory body governing safety (the NRC) requires a minimum level of maintenance.

an unsafe event is $c^s > 0$, of which some fraction θ are borne by the plant, and the remaining fraction $(1-\theta)$ are borne by society.[48]

The firm's optimum is

$$\max_{a} \quad (1 - r(a)) \cdot \pi - r(a) \cdot c^u - c(a, \pi) - s(a) \cdot \theta \cdot c^s \tag{1.4}$$

The social optimum is similar but with $\theta = 1$ (society internalizes all of the safety costs).

The firm's first-order condition is

$$-r'(a) \cdot \pi - r'(a) \cdot c^u - \frac{\partial c(a, \pi)}{\partial a} - s'(a) \cdot \theta \cdot c^s = 0 \tag{1.5}$$

The firm, which does not bear the entire safety cost c^s, exerts less effort a than is socially optimal. However, note that even if the firm internalizes none of the safety costs (i.e., $\theta = 0$), the firm invests in maintenance (because of the reliability costs) that has a positive impact on safety. The social optimum can be achieved if a regulatory agency requires the firm to conduct the optimal level of maintenance. In practice, this may be difficult if the regulatory agency does not have complete information on the cost function $c(a, \pi)$ or the reliability and safety functions $r(a)$ and $s(a)$.

Comparative statics are again straightforward. By the implicit function theorem,

$$\begin{bmatrix} \frac{\partial a}{\partial \pi} \\ \frac{\partial a}{\partial c^u} \\ \frac{\partial a}{\partial \theta} \end{bmatrix} = -\left[-r''(a) \cdot \pi - r''(a) \cdot c^u - \frac{\partial^2 c(a, \pi)}{\partial^2 a} - s'(a) \cdot \theta \cdot c^s\right]^{-1} \cdot \begin{bmatrix} -r'(a) - \frac{\partial^2 c}{\partial a \partial \pi} \\ -r'(a) \\ -s'(a) \cdot c^s \end{bmatrix} \tag{1.6}$$

As before, at the profit maximizing level of a, $\left[-r''(a) \cdot \pi - r''(a) \cdot c^u - \frac{\partial^2 c(a,\pi)}{\partial^2 a} - s'(a) \cdot \theta \cdot c^s\right]$ is negative (by the second order condition, which is satisfied according to the above assumptions).[49] The sign on $\frac{\partial a}{\partial \pi}$ is again indeterminate, and $\frac{\partial a}{\partial c^u}$ is again positive. Since $s'(a) < 0$, $\frac{\partial a}{\partial \theta} > 0$; effort is increasing in the portion θ of the safety cost that the firm internalizes.

At the other extreme, safety could be unrelated to reliability, in that the maintenance effort that lowers the probability of an unplanned outage is separate from any maintenance that improves safety. Denote the maintenance that improves reliability as a^r and the maintenance that improves safety as a^s. Both require expenditures by the plant: $c^r(a^r, \pi)$ and $c^s(a^s, \pi)$, with $c(.) > 0$, $\frac{\partial c(.)}{\partial a} > 0$, and $\frac{\partial^2 c}{\partial a \partial \pi} > 0$ (beyond these assumptions, I make no assumptions on the functional form of $c^r(a^r, \pi)$ as compared to $c^s(a^s, \pi)$). As before, additional maintenance requires a longer time offline, so more revenue is lost (the case where reliability and safety maintenance do not require being offline can also be considered, with $\frac{\partial^2 c}{\partial a \partial \pi} = 0$). The firm's problem is

$$\max_{a^r, a^s} \quad (1 - r(a^r)) \cdot \pi - r(a^r) \cdot c^u - c^r(a^r, \pi) - s(a^s) \cdot \theta \cdot c^s - c^s(a^s, \pi) \tag{1.7}$$

[48] See above for a summary of nuclear reactor liability in the U.S. under the Price-Anderson Act (PAA).

[49] As before, the key assumptions for satisfying the second order condition are that $r''(a) > 0$ and $s''(a) > 0$.

The social optimum is similar but with $\theta = 1$ (society internalizes all of the safety costs).

The firm's first-order conditions are

$$-r'(a^r) \cdot \pi - r'(a^r) \cdot c^u - \frac{\partial c^r(a^r, \pi)}{\partial a^r} = 0 \tag{1.8}$$

$$-s'(a^s) \cdot \theta \cdot c^s - \frac{\partial c^s(a^s, \pi)}{\partial a^s} = 0 \tag{1.9}$$

The firm, like the social planner, equates the marginal cost and benefit of reliability maintenance, so that the firm's choice of a^r is equivalent to the social optimum. However, the firm internalizes only a fraction θ of the benefits associated with improved safety, and exerts a sub-optimal level of effort on safety maintenance. (With perfect information and regulatory oversight, the social optimum could again be achieved through regulation of maintenance levels.) The second order conditions are again satisfied; the Hessian matrix is:

$$\begin{bmatrix} -r''(a^r) \cdot \pi - r''(a^r) \cdot c^u - \frac{\partial^2 c^r(a^r,\pi)}{\partial^2 a^r} & 0 \\ 0 & -s''(a^s) \cdot \theta \cdot c^s - \frac{\partial^2 c^s(a^s,\pi)}{\partial^2 a^s} \end{bmatrix} \tag{1.10}$$

The two diagonal terms are negative, so the matrix is negative definite.

Comparative statics for the firm are:

$$\begin{bmatrix} \frac{\partial a^r}{\partial \pi} & \frac{\partial a^r}{\partial c^u} & \frac{\partial a^r}{\partial \theta} \\ \frac{\partial a^s}{\partial \pi} & \frac{\partial a^s}{\partial c^u} & \frac{\partial a^s}{\partial \theta} \end{bmatrix} = -Hessian^{-1} \cdot \begin{bmatrix} \frac{\partial FOC_1}{\partial \pi} & \frac{\partial FOC_1}{\partial c^u} & \frac{\partial FOC_1}{\partial \theta} \\ \frac{\partial FOC_2}{\partial \pi} & \frac{\partial FOC_2}{\partial c^u} & \frac{\partial FOC_2}{\partial \theta} \end{bmatrix} \tag{1.11}$$

$$= -\begin{bmatrix} -r''(a^r) \cdot \pi - r''(a^r) \cdot c^u - \frac{\partial^2 c^r(a^r,\pi)}{\partial^2 a^r} & 0 \\ 0 & -s''(a^s) \cdot c^s - \frac{\partial^2 c^s(a^s,\pi)}{\partial^2 a^s} \end{bmatrix}^{-1}$$

$$\cdot \begin{bmatrix} -r'(a^r) - \frac{\partial^2 c^r}{\partial a^r \partial \pi} & -r'(a^r) & 0 \\ -\frac{\partial^2 c^s}{\partial a^s \partial \pi} & 0 & -s'(a^s) \cdot c^s \end{bmatrix} \tag{1.12}$$

Denote the above as follows, where $a < 0$, $b < 0$, the sign of c is indeterminate, $d > 0$, $e < 0$, and $f > 0$:

$$= -\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}^{-1} \begin{bmatrix} c & d & 0 \\ e & 0 & f \end{bmatrix} \tag{1.13}$$

$$= -\begin{bmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{bmatrix} \begin{bmatrix} c & d & 0 \\ e & 0 & f \end{bmatrix} \tag{1.14}$$

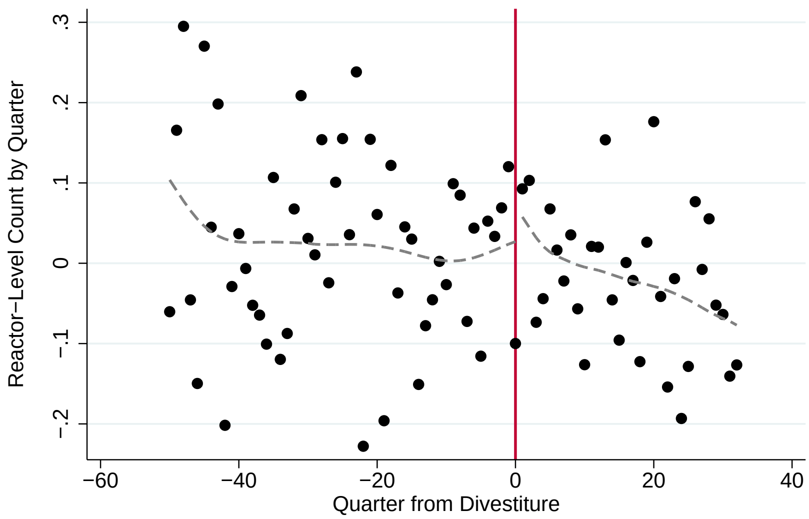
$$= -\begin{bmatrix} \frac{c}{a} & \frac{d}{a} & 0 \\ \frac{e}{b} & 0 & \frac{f}{b} \end{bmatrix} \tag{1.15}$$

$$= \begin{bmatrix} ind & + & 0 \\ - & 0 & + \end{bmatrix} \tag{1.16}$$

Thus $\frac{\partial a^r}{\partial \pi}$ again has an indeterminate sign, $\frac{\partial a^r}{\partial c^u}$ is again positive, and $\frac{\partial a^s}{\partial \theta}$ is again positive. As expected, $\frac{\partial a^r}{\partial \theta}$ and $\frac{\partial a^s}{\partial c^u}$ are both zero: reliability maintenance does not depend on the costs of safety events and vice-versa. Note that $\frac{\partial a^s}{\partial \pi}$ is negative: potential operating profits unambiguously lower the optimal expenditures on safety maintenance. This follows from the assumption that safety maintenance requires that the plant be offline; if we instead assume $\frac{\partial^2 c(a^s,\pi)}{\partial a^s \partial \pi} = 0$, then potential operating profits will not affect the optimal expenditures on safety maintenance.

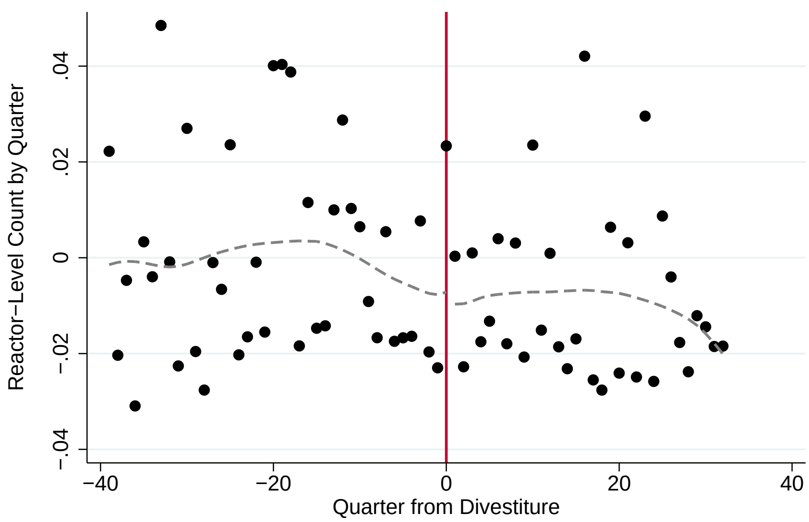
Appendix B: Additional Figures

Figure A.1: Effect of Divestiture on Initiating Events, Quarterly Event Study



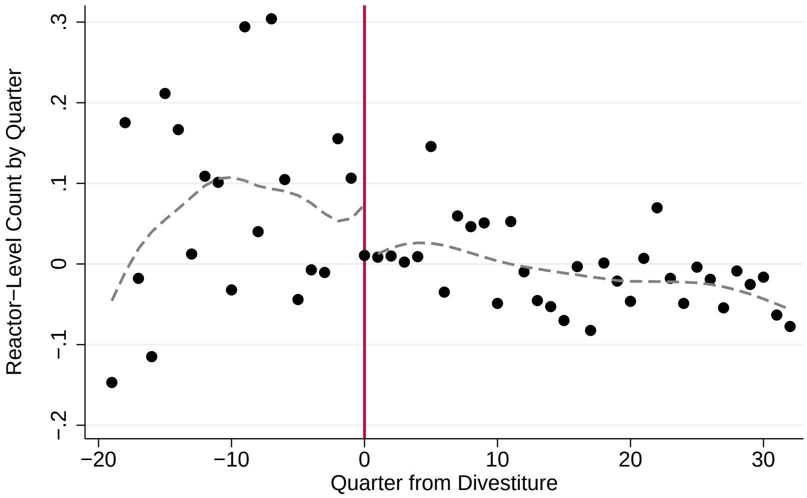
Notes: This figure plots initiating events at divested units relative to non−divested units. Time is normalized relative to divestiture. The median divestiture is in 2001. Quarter−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Figure A.2: Effect of Divestiture on Fires, Quarterly Event Study



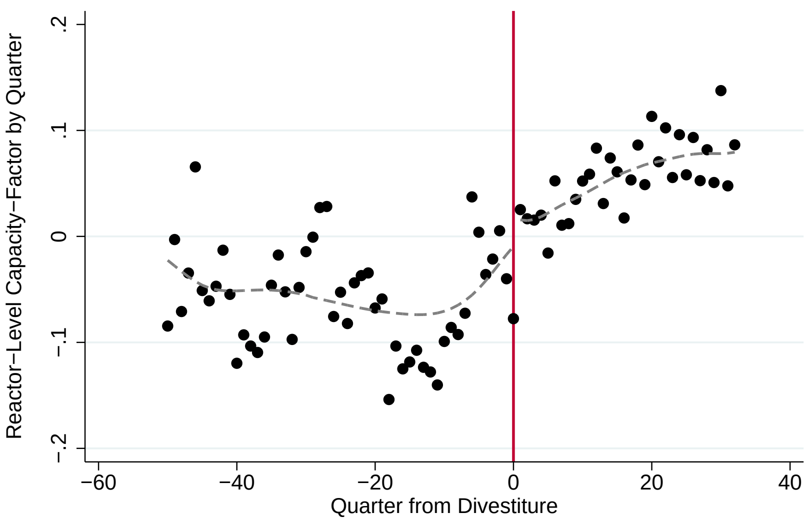
Notes: This figure plots fires at divested units relative to non−divested units. Time is normalized relative to divestiture. The median divestiture is in 2001. Quarter−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Figure A.3: Effect of Divestiture on Escalated Enforcement, Quarterly Event Study



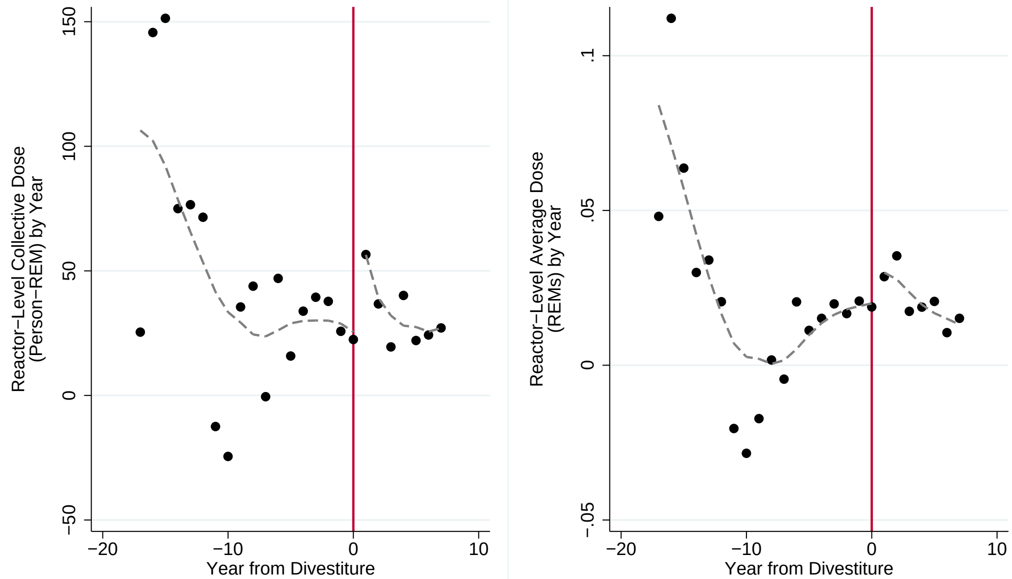
Notes: This figure plots escalated enforcement at divested units relative to non−divested units. Time is normalized relative to divestiture. The median divestiture is in 2001. Quarter−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Figure A.4: Effect of Divestiture on Generation, Quarterly Event Study



Notes: This figure plots capacity factor at divested units relative to non−divested units. Time is normalized relative to divestiture. The median divestiture is in 2001. Quarter−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Figure A.5: Effect of Divestiture on Worker Radiation Exposure, Annual Event Study



Notes: This figure plots collective worker radiation exposure (person−rems) at divested units relative to non−divested units. Time is normalized relative to divestiture. The median divestiture is in 2001. Year−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Notes: This figure plots average worker radiation exposure (rems) at divested units relative to non−divested units. Time is normalized relative to divestiture. The median divestiture is in 2001. Year−of−sample effects have been removed. Dashed grey lines show a lowess smoother in the pre−period and post−period.

Appendix C: Additional Tables

Appendix Table 1: Additional Outcome Variables

	(1)	(2)
Dependent variable:	Withholding Information	Retaliation
A: Dependent variable is not normalized		
Divestiture	-0.460	-1.436
	(1.167)	(0.993)
Specification	Neg Bin	Neg Bin
Year effects	Yes	Yes
Reactor effects	Yes	Yes
Plant effects	No	No
Number of observations	1442	1442
B: Dependent variable is normalized by capacity factor		
Divestiture	-0.658	-1.570
	(1.216)	(1.117)
Specification	Neg Bin	Neg Bin
Year effects	Yes	Yes
Reactor effects	Yes	Yes
Plant effects	No	No
Number of observations	1425	1425

Notes: 81 reactors had no withholding information violations, and 66 reactors had no worker retaliation violations. Data are from the NRC escalated enforcement actions dataset. Withholding Information violations are escalated enforcement actions whose short description refers to "failure to provide information," "withholding information," "violation of 10 CFR 50.9," "lack of complete and accurate information," etc. Worker Retaliation violations are from actions refering to "safety culture," "harassment," "retaliation," "SWCE," "violation of 10 CFR 50.7," etc.

Appendix Table 2: Additional Characteristics for Comparing Divested and Non-Divested Reactors

	never divested	later divested	t-stat	p-value
C. Reactor characteristics:				
Percent PWR	0.78	0.54	1.99	0.05
Age in 1998	18.36	18.80	-0.27	0.79
Capacity (MWe)	959.67	921.92	0.67	0.50
Number of operating reactors at plant	1.87	1.71	1.11	0.27
Manufacturer:				
Babcock & Wilcox	0.09	0.04	0.78	0.44
Combustion Engineering	0.18	0.08	1.15	0.25
General Electric	0.22	0.46	-1.99	0.05
Westinghouse	0.51	0.42	0.71	0.48
Location:				
West	0.15	0.00	2.13	0.04
Midwest	0.18	0.38	-1.68	0.10
South	0.67	0.13	5.02	<0.01
Northeast	0.00	0.50	-5.52	<0.01
D. Maximum generating capacity:				
Licensed capacity	101.94	101.23	1.30	0.20

Notes: For maximum generating capacity (measured annually at reactors), N = 165 for never divested units, 144 for later divested units. For the fixed characteristics, N = 55 for never divested plants, 48 for later divested plants. One reactor (Watts Bar 1) starts commercial operation during this time. T-tests are clustered at the plant level.

Appendix Table 3: Dropping if Capacity Factor <0.01

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforcement	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
Divestiture	-0.205	-0.635	-0.419*	-32.8	-0.024
	(0.127)	(0.431)	(0.252)	(64.6)	(0.022)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2207	1925	1425	1729	1729

Notes: Same regressions as in panel A of table 3, except observations with capacity factor < 0.01 have been dropped (so the sample is the same as for panel B of table 3).

Appendix Table 4: Cumulative Generation

Dependent variable:	(1) Initiating Events	(2) Fires	(3) Escalated Enforcement	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
Divestiture	-0.234*	-0.622	-0.417*	-52.7	-0.022
	(0.123)	(0.431)	(0.249)	(67.7)	(0.020)
Cumulative generation, million MWh	-0.011***	-0.0009	-0.020**	-0.857	0.0003
	(0.003)	(0.011)	(0.008)	(0.568)	(0.0002)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2232	1948	1441	1686	1686
B: Dependent variable is normalized by capacity factor					
Divestiture	-0.362***	-0.766*	-0.538*	-226.1	-0.142
	(0.118)	(0.440)	(0.281)	(278.5)	(0.096)
Cumulative generation, million MWh	-0.008***	0.001	-0.017*	2.261	0.002*
	(0.003)	(0.011)	(0.009)	(2.957)	(0.001)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2194	1923	1424	1666	1666

Appendix Table 5: Lagged Capacity Factor

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforcement	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
	A: Dependent variable is not normalized				
Divestiture	-0.204*	-0.523	-0.213	-38.3	-0.030
	(0.123)	(0.462)	(0.245)	(72.3)	(0.021)
1st lag: Capacity factor	0.692***	0.527	-0.678	138.0	0.149***
	(0.167)	(0.674)	(0.521)	(127.1)	(0.033)
2nd lag: Capacity factor	-0.212	-0.599	-1.074***	-331.7***	-0.021
	(0.148)	(0.671)	(0.274)	(111.0)	(0.034)
3rd lag: Capacity factor	-0.097	-2.256***	0.003	-146.1*	0.006
	(0.192)	(0.568)	(0.287)	(83.4)	(0.028)
Number of observations	2093	1866	1393	1508	1508
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
	B: Dependent variable is normalized by capacity factor				
Divestiture	-0.311**	-0.671	-0.338	-36.1	-0.086
	(0.122)	(0.471)	(0.245)	(258.0)	(0.091)
1st lag: Capacity factor	0.512***	0.581	0.022	-897.2	-0.167
	(0.181)	(0.729)	(0.532)	(1098.1)	(0.529)
2nd lag: Capacity factor	-0.401***	-0.722	-1.287***	-2801.3**	-0.854**
	(0.141)	(0.718)	(0.266)	(1069.0)	(0.322)
3rd lag: Capacity factor	-0.166	2.249***	-0.085	-540.1	-0.113
	(0.187)	(0.564)	(0.286)	(461.0)	(0.162)
Number of observations	2074	1851	1382	1493	1493
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes

Appendix Table 6: Jackknife Regressions

A. Jackknife Regressions, by Plant

A1. Dependent variable not normalized

Variable	Obs	Mean	Min	Max
initiating events	66	-0.192	-0.225	-0.139
fires	66	-0.622	-0.760	-0.494
escalated enforcement	66	-0.426	-0.491	-0.312
collective worker radiation exposure	66	-42.2	-68.4	-10.7
average worker radiation exposure	66	-0.025	-0.032	-0.014

A2. Dependent variable normalized

Variable	Obs	Mean	Min	Max
initiating events	66	-0.335	-0.366	-0.286
fires	66	-0.767	-0.912	-0.644
escalated enforcement	66	-0.552	-0.634	-0.399
collective worker radiation exposure	66	-180.2	-274.9	35.8
average worker radiation exposure	66	-0.108	-0.151	-0.028

B. Jackknife Regressions, by Year

B1. Dependent variable not normalized

Variable	Obs	Mean	Min	Max
initiating events	22	-0.192	-0.233	-0.102
fires	19	-0.623	-0.864	-0.474
escalated enforcement	14	-0.426	-0.595	-0.327
collective worker radiation exposure	35	-42.3	-51.5	-23.6
average worker radiation exposure	35	-0.025	-0.029	-0.021

B2. Dependent variable normalized

Variable	Obs	Mean	Min	Max
initiating events	22	-0.335	-0.378	-0.241
fires	19	-0.768	-1.023	-0.611
escalated enforcement	14	-0.552	-0.721	-0.421
collective worker radiation exposure	35	-180.3	-276.4	-57.5
average worker radiation exposure	35	-0.108	-0.144	-0.053

Notes: Same regressions as in table 3, except a jackknife procedure has been performed. For panel A, the jackknife was by plant. For panel B, the jackknife was by year.

Appendix Table 7: Dropping Exelon

Dependent variable:	(1) Initiating Events	(2) Fires	(3) Escalated Enforcement	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
	A: Dependent variable is not normalized				
Divestiture	-0.217	-0.396	-0.395	32.2	-0.006
	(0.162)	(0.527)	(0.315)	(73.5)	(0.024)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	1873	1627	1204	1534	1534
	B: Dependent variable is normalized by capacity factor				
Divestiture	-0.351**	-0.512	-0.524	190.6	-0.00003
	(0.155)	(0.520)	(0.371)	(178.2)	(0.062)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	1842	1607	1192	1518	1518

Notes: Same regression as in table 3, except dropping all seventeen reactors eventually acquired by Exelon.

Appendix Table 8: Heterogeneity by Reactor Characteristics

Dependent variable:	(1) Initiating Events	(2) Fires	(3) Escalated Enforcement	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
Dependent variable is normalized by capacity factor					
Divestiture, BWR	-0.29**	-0.71	-0.55*	-523.3	-0.185
	(0.14)	(0.53)	(0.33)	(425.6)	(0.180)
Divestiture, PWR	-0.39**	-0.87	-0.56	236.9	-0.014
	(0.18)	(0.59)	(0.37)	(157.9)	(0.069)
Chi-squared stat	0.23	0.05	<0.01	3.50*	0.74
Divestiture, older reactors	-0.21	-0.46	-0.37	-320.3	-0.172
	(0.14)	(0.44)	(0.29)	(385.8)	(0.143)
Divestiture, newer reactors	-0.48***	-1.27*	-0.86**	40.7	-0.007
	(0.17)	(0.71)	(0.37)	(134.4)	(0.060)
Chi-squared stat	2.08	1.27	2.25	1.35	1.89
Divestiture, small reactors	-0.17	-0.80	-0.46	-306.7	-0.211
	(0.15)	(0.51)	(0.28)	(435.4)	(0.167)
Divestiture, large reactors	-0.51***	-0.72	-0.68*	-45.0	0.003
	(0.16)	(0.63)	(0.38)	(216.9)	(0.059)
Chi-squared stat	3.25*	0.01	0.40	0.38	1.92
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2207	1925	1425	1729	1729

Notes: Same regression as table 5 in the paper, except the dependent variable is normalized by capacity factor.

Appendix Table 9: Robustness to Endogenous Timing:
Cutting Window off at 4-Years Post-treatment, with Learning

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforcement	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
Divestiture	0.01	-0.31	-0.16	-63.2	0.026
	(0.21)	(0.92)	(0.43)	(46.3)	(0.027)
Linear trend pre-divestiture	-0.004	-0.07	0.33**	17.6	-0.012
	(0.04)	(0.17)	(0.15)	(24.4)	(0.011)
Linear trend post-divestiture	-0.14	-0.02	-0.73*	-6.5	-0.010
	(0.19)	(0.97)	(0.41)	(25.5)	(0.011)
Specification	Poisson	Poisson	Poisson	OLS	OLS
Year effects	Y	Y	Y	Y	Y
Reactor effects	Y	Y	Y	N	N
Plant effects	N	N	N	Y	Y
Number of observations	2084	1789	1281	1677	1677
B. Dependent variable is normalized by capacity factor					
Divestiture	-0.11	-0.40	-0.15	-46.9	0.248
	(0.19)	(0.93)	(0.47)	(227.5)	(0.244)
Linear trend pre-divestiture	0.02	-0.08	0.25	-23.8	-0.119
	(0.04)	(0.17)	(0.16)	(100.3)	(0.089)
Linear trend post-divestiture	-0.20	-0.05	-0.81*	0.857	0.022
	(0.19)	(0.96)	(0.42)	(112.3)	(0.069)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2046	1764	1264	1657	1657

Notes: Same regressions as in table 7, except dropping observations after four years of divestiture. It is necessary to comapre to table 7 rather than table 3 (main results), because the treatment effect changes over time.

Appendix Table 10: Deregulation Dates

Dependent variable:	(1) Initiating Events	(2) Fires	(3) Escalated Enforce-ment	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
dereg_main	-0.006	-0.651*	0.017	-2.82	-0.005
	(0.122)	(0.377)	(0.249)	(77.9)	(0.028)
dereg_law	0.009	-0.480	-0.114	-69.9	-0.008
	(0.126)	(0.350)	(0.252)	(70.6)	(0.025)
dereg_retail	-0.103	-0.414	-0.286	-45.7	0.009
	(0.133)	(0.415)	(0.272)	(66.8)	(0.023)
dereg_implement	0.020	-0.352	-0.298	-58.0	-0.004
	(0.108)	(0.348)	(0.268)	(63.4)	(0.023)
B: Dependent variable is normalized by capacity factor					
dereg_main	-0.123	-0.781**	-0.019	-163.9	-0.196
	(0.121)	(0.385)	(0.282)	(290.5)	(0.120)
dereg_law	-0.094	-0.576	-0.159	-245.6	-0.162
	(0.124)	(0.360)	(0.283)	(263.0)	(0.102)
dereg_retail	-0.228*	-0.502	-0.427	-302.9	-0.192
	(0.129)	(0.416)	(0.302)	(279.6)	(0.121)
dereg_implement	-0.094	-0.444	-0.437	-268.2	-0.173*
	(0.102)	(0.357)	(0.296)	(256.4)	(0.103)

Notes: Each dependent variable is for a separate regression (eight regressions total). Regressions are otherwise the same as in table 3. dereg_main: this variable turns on when legislation is first passed, but only in states where activities were never suspended. dereg_law turns on when legislation is first passed and turns off with when activities are suspended. dereg_retail turns on when retail choice begins and turns off when activities are suspended. dereg_implement turns on in the year Craig and Savage (2009) use for implementation, turns off when activities are suspended. Legislation, retail choice, and suspension dates are taken from the EIA's "Status of State Electric Industry Restructuring Activity" (February 2003), accessed October 2011 at http://www.eia.gov/cneaf/electricity/chg_str/restructure.pdf.

Appendix Table 11: Intra-Firm Spillovers

Dependent variable:	(1) Initiating Events	(2) Fires	(3) Escalated Enforcement	(4) Collective Worker Radiation Exposure (person-rems)	(5) Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
Divestiture	-0.22	-0.60	-0.56**	-72.4	-0.024
	(0.13)	(0.47)	(0.23)	(63.8)	(0.022)
Co-owned	-0.05	0.02	-0.80*	-144.3	-0.001
	(0.18)	(0.45)	(0.46)	(121.6)	(0.029)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2245	1950	1442	1749	1749
B: Dependent variable is normalized by capacity factor					
Divestiture	-0.36***	-0.74	-0.68***	-187.9	-0.067
	(0.13)	(0.48)	(0.25)	(276.5)	(0.115)
Co-owned	-0.01	0.06	-0.72	-63.6	0.129
	(0.17)	(0.45)	(0.48)	(282.8)	(0.097)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2207	1925	1425	1729	1729

Notes: Co-owned is a dummy equal to 1 if the reactor is not divested, but is owned by a company operating divested units (Dominion, Entergy, and NextEra). Thus the omitted group is non-divested reactors whose parent company operates no divested reactors.

Appendix Table 12: Consolidation

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforce-ment	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
A: Dependent variable is not normalized					
Divestiture	-0.29*	-0.21	-0.40	73.4	0.0002
	(0.16)	(0.56)	(0.33)	(68.4)	(0.027)
Consolidation	0.02	-0.07	-0.004	-23.4**	-0.005
	(0.02)	(0.06)	(0.03)	(10.9)	(0.005)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2245	1950	1442	1749	1749
B: Dependent variable is normalized by capacity factor					
Divestiture	-0.41***	-0.32	-0.54	203.3	-0.071
	(0.16)	(0.57)	(0.39)	(201.0)	(0.095)
Consolidation	0.01	-0.07	-0.001	-78.0	-0.007
	(0.02)	(0.07)	(0.04)	(53.0)	(0.030)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2207	1925	1425	1729	1729

Notes: Consolidation is a count variable, equal to the number of other reactors owned by the parent company.

Appendix Table 13: Event Study

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Initiating Events	Fires	Escalated Enforce-ment	Collective Worker Radiation Exposure (person-rems)	Average Worker Radiation Exposure (rems)
			A: Dependent variable is not normalized		
>=5 years pre-divestiture	-0.018	0.561	-0.484	-9.94	0.023
	(0.156)	(0.792)	(0.444)	(76.46)	(0.028)
4 years pre-divestiture	-0.245	0.890	0.171	14.97	0.007
	(0.236)	(0.910)	(0.391)	(34.29)	(0.017)
3 years pre-divestiture	-0.219	0.957	-0.014	23.14	0.007
	(0.232)	(0.883)	(0.396)	(22.80)	(0.012)
2 years pre-divestiture	-0.020	-16.475***	0.443	16.83	0.004
	(0.198)	(0.746)	(0.391)	(45.37)	(0.019)
1 year pre-divestiture	normalized to zero	normalized to zero	normalized to zero	normalized to zero	normalized to zero
divestiture	0.058	0.318	0.302	-11.58	0.002
	(0.216)	(0.926)	(0.390)	(41.14)	(0.016)
1 year post-divestiture	-0.131	-0.035	-0.151	40.63	0.003
	(0.261)	(1.045)	(0.482)	(35.71)	(0.013)
2 years post-divestiture	-0.262	-0.091	-0.059	-0.76	0.013
	(0.245)	(1.066)	(0.423)	(34.03)	(0.017)
3 years post-divestiture	-0.167	-0.048	-0.672	-34.08	-0.001
	(0.222)	(1.022)	(0.588)	(29.27)	(0.018)
4 years post-divestiture	-0.184	0.184	-0.729	-44.38	-0.018
	(0.259)	(1.048)	(0.570)	(28.71)	(0.018)
>=5 years post-divestiture	-0.422*	-0.362	-0.518	-95.76*	-0.022
	(0.233)	(0.946)	(0.377)	(48.10)	(0.022)
Specification	Neg Bin	Neg Bin	Neg Bin	OLS	OLS
Year effects	Yes	Yes	Yes	Yes	Yes
Reactor effects	Yes	Yes	Yes	No	No
Plant effects	No	No	No	Yes	Yes
Number of observations	2245	1950	1442	1749	1749

Notes: There are no fires 2 years pre-divestiture at any plant.

Chapter 2

Farm Acreage Shocks and Crop Prices: An SVAR Approach to Understanding the Impacts of Biofuels[1]

with Maximilian Auffhammer and Peter Berck

This chapter is published as: Hausman, Catherine, Maximilian Auffhammer, and Peter Berck. 2012. "Farm Acreage Shocks and Crop Prices: An SVAR Approach to Understanding the Impacts of Biofuels." *Environmental and Resource Economics* 53(1): 117-136. The final publication is available at link.springer.com.

2.1 Introduction

Biofuels have been promoted as an alternative to petroleum products that bypasses some of the fundamental problems with the oil market: supporters claim that it is renewable (whereas conventional oil is exhaustible), produced in the U.S. (as opposed to regimes in some cases unfriendly to the U.S.), and carbon-friendly. As biofuels production has expanded, however, concerns have been raised about their direct and indirect impacts, particularly on land use and on food prices. The last ten years have seen tremendous expansion in biofuels production, particularly in corn ethanol in the United States, at the same time that commodity prices (e.g., corn) have experienced significant spikes. In 2006, 2.1 billion bushels of corn went to ethanol production (approximately 15% of all corn production); in 2007 this rose to over 3 billion bushels (20% of corn production). From 2005 to 2006, corn prices rose by 47%, and from 2006 to 2007 they rose by another 28%. Not surprisingly, the "food

[1]We are grateful to the editor and two anonymous reviewers whose suggestions have greatly improved the manuscript. Many thanks to Michael Jansson for helpful suggestions. Financial support provided by the Energy Biosciences Institute. All views are those of the authors alone and do not represent those of the Energy Biosciences Institute.

versus fuel" debate has economists and policy makers asking how much of the increase in corn prices is due to the increased demand for corn from ethanol producers. On October 3, 2007, the BBC argued that "[i]t is one of the most hotly debated environmental topics of the year - whether the drive to produce alternative so-called green fuels will take food from the mouths of the hungry." As Roberts and Schlenker (2010) point out, the economic literature has yet to agree on the magnitude of these hypothesized effects.

While corn for ethanol has, to date, competed directly with the production of food commodities, scientists hope that future so-called "second generation" biofuels will use non-food crops and marginal lands (Heaton et al. 2008, Hill et al. 2006, and Robertson et al. 2008). The primary first-generation biofuel is corn ethanol, which uses a food crop and conventional sugar to ethanol fermentation to produce fuels used in transportation. Second-generation biofuels use non-food crops (e.g., miscanthus or switchgrass) and a different technology, in which cellulosic plant material is converted into ethanol. This paper contributes to a growing literature analyzing the effect of corn ethanol on commodity and food prices. In this paper, we examine econometrically the effects of exogenous shocks in acreage supply on food crop prices, using 50 years of U.S. data. We analyze the effects of both shocks to a crop's own acreage and shocks to total cropland. Our approach allows us to calculate the effect of dedicating existing cropland to biofuels feedstock production as well as the effect of dedicating non-crop lands. We focus on shocks to acreage rather than shocks to crop yields, as there has been no evidence to date that biofuels production will change the yields of crops grown for food. Accordingly, the shock to acreage that we estimate is a useful summary statistic for shocks to the production of crops such as corn and soybeans.[2] These shocks to acreage can be thought of as a shock to the residual supply for corn for food; corn can be dedicated to food uses or fuel uses, so a shock to the demand for fuel corn translates into a shock to the residual supply for food corn.

Our paper makes three main contributions to this literature. First, we rely on econometrically estimated coefficients using observed data for the U.S., an approach that steps away from the previous analyses that assumed supply and demand elasticities (e.g., Sexton et al. (2008) and Rajagopal et al. (2007)). Second, we study the effects of removing either existing cropland or non-crop farmland on food crop prices. Third, in order to address the obvious endogeneity of supply and demand systems, we borrow econometric tools from the macroeconomic literature that leverage the timing of exogenous shocks.

The dynamic nature of agricultural production makes this question ideally suited for a structural vector auto-regression (SVAR) of the sort routinely used in the macroeconomic literature (e.g., Christiano et al. (2005)). This framework allows us to leverage the timing of planting decisions versus harvest outcomes with a classic time-series methodology. We estimate a system of equations to explain the relationship between total cropland, corn and soybean acreage, and corn and soybean spot and futures prices. In particular, we use a factor-augmented structural vector auto-regression to allow for exogenous shocks to the entire system, including supply-side shocks, such as a spike in farm input prices, or a demand-side shock, such as increased foreign demand.

[2]We do allow the shock to acreage to feed through to shocks to yield, since farmers will re-optimize their yield decisions based on acreage decisions.

Time-series models such as SVARs are widely used in macroeconomic applications where variables are jointly determined, adjustment to long-run equilibrium is not instantaneous (which implies the importance of including lags in the model), and the underlying data generating process follows a specific timing mechanism. The above are all characteristics of an agricultural supply model, in which price and acreage are jointly determined and in which the effects of a shock can last several periods.[3] We show that a structural VAR can be used to leverage the sequencing inherent in agricultural supply. The resulting model is able to capture dynamics that may be missed with other models. Also, the system's dynamic nature allows us to estimate forecast error variance decompositions (FEVDs), which explain the percentage of variance that comes from specific shocks, and impulse response functions (IRFs), which trace out the effect of exogenous shocks across time.

Estimation results show that a reduction in corn area of one million acres (approximately 1% of U.S. corn acreage) leads to a corn price increase of $0.04 per bushel (an approximately 1% increase). To put this in perspective, consider that from 2006 to 2007 (a boom production year), corn acreage dedicated to ethanol increased by about 6 million acres in the United States. At the same time, corn prices rose by $0.88 per bushel. Thus our model finds that approximately 27% of the price increase was due to new ethanol production. For a negative shock in soybean acreage, we find a price increase of $0.23 per bushel, with a wider variation across alternative specifications. This much larger magnitude is partly explained by the fact that US soybean acreage has a much larger share of world production than does US corn acreage.

For the scenario in which one million acres is moved from non-crop farmland (e.g., pasture and idle lands) to crop production, while holding corn or soybean acreage constant, we find a price decrease of $0.05 per bushel of corn and $0.08 per bushel of soybeans. The intuition for this result is that, while corn (or soybean) acreage is held constant, the production of substitute crops increases. Accordingly, demand for the crop falls. Historically, this has corresponded to increased production of other grains, but the intuition is consistent with a scenario of increased production of substitute biofuels feedstocks. Thus we find robust results that removing acreage from food production, in order to grow biofuels feedstocks, increases food crop prices. On the other hand, switching non-crop farmland to a substitute crop can lower food crop prices. As the U.S. and E.U. formulate biofuels policies, this empirical evidence on the costs for food production and commodity prices should be taken into account. This paper proceeds as follows: in section 2.2, we summarize the related literature. In section 2.3, we develop the econometric framework; in section 2.4 we present

[3]Previous work on dynamic agricultural systems beyond Nerlove's model (1956) is relatively limited. In the Nerlovian framework, farmers make production decisions (including acreage decisions) according to their expectations on crop prices and input prices. At planting time, they observe only last year's prices, planting-time spot prices, and futures prices for harvest-time delivery. Their decision can be modeled in a partial-adjustment framework, in which acreage at time t is a function of acreage at time $t-1$, plus futures prices and past spot prices. This equation appears in our model, but we generalize the framework by also modeling the movement of prices. Mushtaq and Dawson (2002) use a recursive vector auto-regression approach to investigate acreage response of various crops in Pakistan. They find that this approach is more appropriate than a Nerlovian partial-adjustment model, particularly in explaining adjustments to long-run equilibrium. However since their interest is in acreage response, they do not report the impact of shocks to acreage on prices.

the data. Section 2.5 shows the main results, section 2.6 presents the robustness checks, and in section 2.7 we conclude.

2.2 Related Literature

A number of papers have simulated the impact of biofuels production on various economic outcomes, generally using supply and demand elasticities drawn from the literature. Banse et al. (2008) use a global CGE model (modified from the GTAP model) to analyze the trade impacts of an EU Biofuels Directive. They find that cereals prices actually decline in the long run, but less than they would without the directive. This finding is a combination of an assumed inelastic demand and a high rate of productivity change. Rajagopal et al. (2007) use a stylized partial equilibrium model and find a 21% increase in corn price attributable to a $0.51 ethanol production tax credit in the US in 2006. A 2009 CBO study estimated that 10 to 15% of the food price increase from April 2007 to April 2008 was attributable to expanded ethanol production. The estimated impact on corn prices for the same period is higher: between 50 and 80 cents per bushel, or 28 to 47 percent of the total corn price increase. Rosegrant (2008) uses a partial equilibrium model from IFPRI and finds that biofuel demand accounted for 39% of the corn price increase from 2000 to 2007. Chakravorty et al. (2011) find in simulations that biofuels mandates drive land allocation changes rather than large food price increases. Two other oft-cited papers argue that biofuels policy has driven up corn prices but do not make direct calculations (Abbott, Hurt, and Tyner (2008) and Mitchell (2008)). Finally, Chakravorty, Hubert and Nostbakken (2009) provide an extensive literature review on the various models that have been applied to biofuels, including partial equilibrium agricultural models from FAPRI, IFPRI, and IIASA, and general equilibrium models such as GTAP. Naylor et al. (2007) provide a useful summary of predicted crop price changes under the various biofuels scenarios found in the literature.

In related work, some papers analyze the welfare impacts in developing countries (e.g., Runge and Senauer, 2007). Naylor et al. (2007) focus on the food security impacts of biofuels expansion. Roberts and Schlenker (2010) construct an elegant system of supply and demand for food calories, which they estimate using novel instrumental variable techniques based on weather shocks. The resulting price increases imply a large change in global consumer surplus.

Several papers have looked at some of the other possible explanations for recent crop prices. These explanations include export restrictions, growing food demand from developing countries, low investment and hence low productivity growth, weather shocks, crop diseases, depreciation of the U.S. dollar, increases in the price of crude oil, production cost increases, speculation in commodity markets, and the additional impact of low stocks (e.g., Abbott, Hurt, and Tyner (2008), Chakravorty et al. (2011), Headey and Fan (2008), and Mitchell (2008)).

Additionally, recent papers examine at the impact of biofuels policies on economic outcomes other than crop prices. Ando, Khanna, and Taheripour (2010) evaluate the impact of the Renewable Fuel Standard on the transportation sector; Khanna, Ando, and Taheripour (2008) evaluate the impact of ethanol production on greenhouse gas emissions and congestion externalities.

Finally, the structural vector auto-regression framework has been used in previous studies examining dynamic interactions between corn, ethanol, gasoline, and crude oil markets. Zhang et al. (2007) develop a model of the U.S. fuel market that focuses on the interaction between MTBE and ethanol, two fuel additives. Cha and Bae (2011) employ a structural VAR, identified through sign restrictions, to estimate the impact of shocks to international oil prices and shocks to demand for corn exports on corn prices, ethanol demand for corn, and feed demand for corn. McPhail (2011) estimates a recursive structural VAR model of crude oil, gasoline, and ethanol markets, focusing on the difference between ethanol demand shocks and ethanol supply shocks.

2.3 Econometric Framework

2.3.1 Structural Vector Auto-Regression: Scenario 1

We apply a structural vector auto-regression[4] to analyze what we call scenario 1, in which food acreage is removed and dedicated to a biofuels feedstock.[5] In this framework, we use a system of equations to explain the relationship between corn and soybean acreage, total cropland, and corn and soybean spot and futures price. We impose identification restrictions that take advantage of the timing of planting decisions in the United States. That is, agricultural producers set their acreage at planting time according to their expectation of harvest-time prices. In the classic Nerlovian framework, this implies that current acreage is a function of past acreage, past prices and futures prices, and supply-side variables such as input prices, which are the only variables observable to farmers at planting time. Price at harvest is then a function of production (acreage times yield) and a number of demand-side market forces. Accordingly, the vector of variables of interest is

$$y_t \equiv \begin{bmatrix} corn\, futures_t \\ soy\, futures_t \\ supply\, variables_t \\ corn\, yield_t \\ corn\, acreage_t \\ total\, farmland_t \\ demand\, variables_t \\ corn\, harvest\, price_t \\ soy\, harvest\, price_t \end{bmatrix}$$

A generic vector auto-regression with exogenous variables x_t has the following structure

[4] In time series literature, "structural" vector auto-regressions refer to vector auto-regressions that allow for causal interpretations. This use of the word "structural" is different from that in the general econometrics literature; the system of equations need not explicitly model an optimization problem.

[5] Note that what matters for this model is not which feedstock is grown (e.g., corn versus switchgrass) but where that feedstock is grown (land previously dedicated to food production versus non-crop farmland).

$$Ay_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_k y_{t-k} + Cx_t + B\varepsilon_t$$

where $\varepsilon_t \sim N(0, I_K)$ and $E(\varepsilon_s \varepsilon_t = 0), s \neq t$. We then impose restrictions according to our identifying assumptions. We can re-write the above equation as follows

$$y_t = A^{-1}A_1 y_{t-1} + A^{-1}A_2 y_{t-2} + ... + A^{-1}A_k y_{t-k} + A^{-1}Cx_t + u_t$$

where $u_t = A^{-1}B\varepsilon_t$, implying that u_t follows a white noise process. Our identifying assumptions are

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 1 & 0 & 0 & 0 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & 0 & 1 & 0 & 0 & 0 & 0 \\ a_{61} & a_{62} & a_{63} & 0 & 0 & 1 & 0 & 0 & 0 \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & 1 & 0 & 0 \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & 1 & 0 \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{44} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{66} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_{77} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{88} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{99} \end{bmatrix}$$

Note that there are no restrictions on the lags (A_1through A_k) or on the coefficients on exogenous coefficients. The restrictions on A come from the timing of the agricultural production process in the United States. That is, corn and soybean futures (which are observed in March for harvest-time delivery) and supply-side variables (e.g., input costs and loan rates) are generated before acreage decisions or fall prices have been observed; accordingly, they are functions of only lagged and exogenous variables. Total farmland, corn acreage, and corn yields are determined after futures prices have been observed but before fall prices are known. Thus the information set at time t for these variables consists of futures prices, supply-side variables, and lagged and exogenous variables. This is similar to the acreage function in a Nerlovian (partial-adjustment) framework. Demand-side variables (such as foreign production, affecting demand for US exports) that are determined in the summer are functions of futures prices, supply-side variables, US acreage and yields, and lagged and exogenous variables. Finally, harvest time prices are a function of that year's futures prices, supply-side variables, acreage and yield decisions, and demand-side variables. The structure of B imposes orthogonality of contemporary structural shocks. The exogenous variables x_t are constants and time trends. We focus on the case of one lag ($k = 1$), which we find is preferred according to Schwarz's Bayesian information criteria, but examine the robustness of our estimates to the inclusion of additional lags. The system is estimated via maximum likelihood.

2.3.2 Diffusion Indices

As mentioned above, the orthogonality conditions on the matrix B require that there be no omitted variables. Accordingly we control for spring-time supply-side variables (such as input prices and agricultural loan rates) and summer-time demand-side variables (such as US income and foreign production, which affects demand for US exports). However the curse of dimensionality prevents us from including all of these variables in the system; we would quickly run out of degrees of freedom. Thus we include diffusion indices (also known as principal components or factors) to control for these variables while avoiding the curse of dimensionality inherent in large VAR models.

Stock and Watson (2002) show that a large number of time series variables can be summarized with a few indices using principal components analysis. The end result is a linear combination of the original time series, with the linear coefficients chosen to incorporate as much of the variation in the original series as possible. This nonparametric approach begins with the objective function

$$(\hat{F}, \hat{\Lambda}) = argmin[(NT)^{-1} \sum_i \sum_t (x_{it} - \lambda_i F_t)^2]$$

where (F) are the factors and (Λ) the factor loadings. This is solved by setting $\hat{\Lambda}$ equal to the eigenvectors of $X'X$ corresponding to the largest eigenvalues. $\hat{F}$ is then found by setting $\hat{F} = X'\hat{\Lambda}/N$. This approach is applied separately to two sets of time series variables, one consisting of variables affecting crop supply and one of variables affecting crop demand. For each variable used in the indices we test for a unit root, take the first difference of the natural log, and standardize to mean zero and unit variance. Then we estimate a supply-side diffusion index and a demand-side diffusion index, which are both incorporated into the structural vector auto-regression. Thus we are able to control for international and domestic macroeconomic disturbances.

2.3.3 Robustness Checks

One potential concern with the above factor-augmented SVAR is its large size. Generally, smaller systems perform better in this framework than do larger systems. This concern obviously needs to be balanced with the potential of omitted variables bias. To address the concern, we also estimate a sparser model as a robustness check. This model contains only corn and soybean futures prices, corn acreage, farmland, and corn and soybean farmgate prices. Accordingly, the identifying restrictions on matrices A and B are:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 & 0 & 0 \\ a_{41} & a_{42} & 0 & 1 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & 1 & 0 \\ a_{61} & a_{62} & a_{63} & a_{64} & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{66} \end{bmatrix}$$

Additional robustness checks include using a log/log specification, allowing additional lags, varying the time period studied, and varying the factor indices for supply- and demand-side variables.

2.3.4 SVAR Framework for Scenario 2

In scenario 2, we consider the effect of growing a biofuels feedstock on acreage not previously dedicated to food-crop production. We hold own (corn or soybean) acreage constant, while decreasing non-crop farmland (thus increasing total cropland). In a robustness check, we show that results are fairly similar if the system is estimated with a positive shock to total cropland rather than a negative shock to non-crop farmland. Accordingly, the vector of variables of interest is

$$y_t \equiv \begin{bmatrix} corn\,futures_t \\ soy\,futures_t \\ supply\,variables_t \\ corn\,yield_t \\ corn\,acreage_t \\ non-crop\,farmland_t \\ demand\,variables_t \\ corn\,harvest\,price_t \\ soy\,harvest\,price_t \end{bmatrix}$$

The identifying restrictions on A and B are the same as for scenario 1. We again consider various robustness checks, including a sparser specification (ignoring yields, supply variables, and demand variables), additional lags, a log/log functional form, different supply- and demand-side indexes, and a varying time frame.

2.3.5 Forecast Error Variance Decomposition and Impulse Response Functions

The dynamics of the system imply that interpretation of either the reduced form or structural coefficients is difficult. Two tools for analyzing the coefficients are forecast error variance decompositions (FEVDs) and impulse response functions (IRFs). Forecast error variance decomposition tells us the percentage of the forecasting error for a variable due to a specific shock at a given horizon. Following Lutkepohl (1993), we define the FEVD at horizon h as $\omega_{jk,h} = \frac{\psi_{jk,0}^2 + ... + \psi_{jk,h-1}^2}{MSE[\hat{y}_{j,t}(h)]}$, where $\psi_{mn,i}$ denotes the mn-th element of $(A^{-1}A_1)^i A^{-1}B$, and $MSE[\hat{y}_{j,t}(h)] = E(y_{j,t+h} - \hat{y}_{j,t}(h))^2 = \sum_{k=1}^{K}(\psi_{jk,0}^2 + ... + \psi_{jk,h-1}^2)$. Thus the FEVD at horizon h (for instance, $h = 2$) estimates the percentage of the total forecast error that comes from each orthogonalized structural shock.

The dynamic nature of the above system also allows us to estimate impulse response functions (IRFs), which trace out the effect of exogenous shocks on realizations of the random variables across time. Working from the VAR's moving average representation, we can write the structural impulse response function as follows: $y_t = \mu + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$. Thus the structural impulse response function traces each element of Ψ_i for each time period following a shock in period $i = 0$.

We report confidence intervals based on the delta method, as bootstrapped confidence intervals in our over-identified system require a great deal of computing power. Clearly this is an imperfect solution, given how the delta method can perform in a highly non-linear system. We do calculate confidence intervals for the main specifications using a parametric bootstrap method, and we find that the intervals are quite similar to those computed with the delta method. The bootstrapped intervals are slightly narrower but do not change the inference.

Finally, we can compute cumulative impulse response functions from the coefficients. The above structural impulse response functions give the dynamic path of each variable following a shock in period $i = 0$. A cumulative impulse response function gives the dynamic path of each variable as the shock is repeated in each period $i = 0, 1, 2...n$. and is given by $\Xi_n = \sum_{i=0}^{n} \Psi_i$.

2.4 Data

Data are obtained for US production of corn and soybeans from 1956 to 2007. Data on farmland, planted corn acreage, and planted soybean acreage (all measured in thousand acres) are obtained from the National Agricultural Statistics Service (NASS) at the USDA. Crop prices paid to farmers, in dollars per bushel, are also obtained from NASS. These are then deflated by the third-quarter GDP deflator, obtained from the Bureau of Economic Analysis (BEA), into 2007 dollars per bushel. Corn and soybean futures, available from Datastream and the Wall Street Journal, are planting-time quotes for delivery at harvest time. For corn, they are the March 31 closing price for delivery in September. For soybeans, they are the March 31 closing price for delivery in November. The futures prices are deflated by the first-quarter GDP deflator (from the BEA). Corn and soybean yields, in bushels per planted acre,[6] are calculated from NASS production and acreage data. Data on total cropland (also measured in thousand acres) is obtained from NASS, but unfortunately it is only available to 2006. Accordingly, as a robustness check, we also use NASS data on acreage devoted to principal crops.[7]

All variables are examined for evidence of unit roots (table 1). We consider augmented Dickey-Fuller unit root tests and Phillips-Perron tests, with and without trends, for all variables. A unit root is rejected at the 5% level for corn acreage, corn and soy yields, and the supply and demand diffusion indices. We also perform a Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for each variable, with and without trends. Trend stationarity is rejected at the 5% level for corn and soy spot and futures, soy acreage, farmland, non-crop farmland, and soy yields.[8]

For the supply-side diffusion index, eleven variables are used. As these affect farmer planting decisions, they reflect only information available up until planting time (for corn and soybeans, first quarter data). Oil prices are given by West Texas Intermediate Oil Prices at the end of the first quarter, available from Global Financial Data (GFD). The national average loan rate for corn, soybeans, and wheat is available from the Commodity Research Bureau (CRB) from 1956 to 2003, and from the Economic Research Service (ERS) at the USDA for 2004 to 2007. Input prices, including automobiles, two USDA-calculated indices of producer prices paid, building materials, and farm wages are available from NASS. All prices are deflated by the BEA's GDP deflator. Note that fertilizer prices, an important input affecting crop profitability, are incorporated in the USDA-calculated indices of prices paid.

For the demand-side diffusion index, data should reflect information available up until (and including) harvest time. For corn and soybeans, this implies fourth quarter data. Oil prices are given by West Texas Intermediate Oil Prices, available from Global Financial Data (GFD). Third-quarter US Gross National Product is obtained from GFD. Oil prices and GNP are deflated by the US GDP deflator, also from GFD. Corn and soybean production from Southern Hemisphere countries (Argentina and Brazil), which harvest during the US summer months, are available from the Food and Agriculture Organization of the United Nations (FAO). The ideal diffusion index would also incorporate GDP from importing countries, but this is not reliably available on a quarterly basis. Since the timing of innovations is important for the ordering of the VAR, incorporating GDP data updated yearly would be inappropriate.

The estimated indices are linear combinations of the logged, differenced, standardized variables. The first supply index accounts for 30% of the variation in all the series. The two USDA-computed input prices series and oil prices have the largest coefficients in this

[6]Results are extremely similar if we use yields from harvested, rather than planted, acres (figure 3).

[7]Principal crops include barley, beans, beets, corn, cotton, flax, hay, oats, peanuts, potatoes, rice, rye, sorghum, soy, sugarcane, tobacco, and wheat. For most crops, data on planted acreage is used; for beans, sugarcane, and tobacco only data on harvested acreage is available.

[8]We also test for Granger causality between all variables. As expected, lagged values of corn acreage and corn and soy spot prices help predict current values of corn futures prices; soy spot prices help predict soy futures prices; etc. Full results are available upon request.

linear combination, thus the index loads primarily onto them. The second supply index accounts for 20% of the variation and loads primarily onto loan prices. The first demand index accounts for 24% of the variation in all the series and loads onto all six series. The second demand index accounts for 23% of the variation in all the series and loads primarily onto soy and corn production in Argentina.[9]

2.5 Results

We estimate the parameters using maximum likelihood,[10] then compute the percentage variance due to shocks (the Forecast Error Variance Decompositions) and the dynamic path of the variables following a shock (the Impulse Response Functions). The one-lag nine-equation SVAR gives the structural FEVD for scenarios 1 and 2 with corn acreage shown in tables 2.1 and 2.2.[11] Shocks to corn acreage explain 7% of the one-step ahead forecast error in corn harvest prices, for scenario 1. The supply index contributes to 16% of the error, and unobserved variation to 51%. For scenario 2, shocks to non-crop farmland explain 4% of the one-step ahead forecast error in corn harvest prices. The majority of the variation comes from the supply index (21%) and unobserved variables (58%).[12]

The structural FEVDs for the soy scenarios are shown in tables 2.3 and 2.4. For scenario 1, shocks to soybean acreage contribute to slightly more (9%) of the own-crop harvest price than was the case for corn. The majority of the forecast error continues to come from variation in the supply index (24%) and unobserved variables (53%). The largest contributors to the forecast error are again the supply index and unobserved variables. Thus for both crops and both scenarios, shocks to acreage have historically contributed to a fairly small percentage of the one-step forecast error in spot prices.

Figure 1 shows selected structural impulse response functions estimated from the one-lag nine-equation model for scenarios 1 and 2. In particular, we show the IRF graphs for the effect of negative corn (and soybean) acreage shocks to own price and the effect of negative non-crop farmland shocks to own corn and soybean prices.[13] For every 1 million acres of corn production removed, corn price increases in the first period by \$0.04 per bushel.[14] The effect lasts one additional period, and then falls back to zero. For every 1 million acres of soy production removed, soy price increases in the first period by \$0.23 per bushel. The effect is

[9]Detailed descriptions of the estimation of these indices are available upon request.

[10]A table of estimated coefficients is available upon request.

[11]Standard errors are not shown because of space considerations. They are available upon request.

[12]FEVDs at longer horizons are available upon request. For the two-step forecast error in corn harvest prices, corn acreage still explains 7%. At two periods, shocks to non-crop farmland explain 11% of the forecast error. At two periods, shocks to soybean acreage contribute 7% of the forecast error in harvest price. For scenario 2, shocks to non-crop farmland contribute to 13% of the two-step soybean price forecast error. Thus for both crops and both scenarios, shocks to acreage have contributed to a fairly small percentage of the two-step forecast error in spot prices.

[13]Each SVAR has nine equations and therefore 81 estimated IRFs. These results are available upon request.

[14]All results reported are normalized to shocks of 1 million acres. The IRF as described in the methodology section yields estimates of a change in price following a shock of size ε in the acreage equation. This response can be rescaled by the corresponding element of matrix B to give a change in price following a shock of size 1 million acres.

more persistent than it was for corn, although it does fall back towards zero. We expect the effect to be higher for soybeans than for corn, as US production of soybeans has historically had a much larger share of world production than has US corn.

For a negative one million acre shock to non-crop farmland (holding corn acreage constant), corn price falls. The initial effect is \$0.02 per bushel, peaking at \$0.05 per bushel. Eventually the effect returns to zero. Holding soybean acreage constant, the negative shock to non-crop farmland leads to a soybean price decrease of \$0.08 per bushel. What appears to be happening for both crops is that, holding own acreage constant, acreage of other crops is rising. Since grains are largely substitutable, this takes pressure off of the demand for the crop. Since own acreage was held constant, supply doesn't change, and the fall in demand lowers prices. This intuition could correspond historically to increases in, for instance, other food crops like wheat. There is no reason to expect the story not to hold for other biofuels feedstocks, e.g., miscanthus or switchgrass.

Additionally, we compute cumulative impulse response functions (CIRFs). While the structural IRFs described above show the dynamic path of prices following a one-time acreage shock, a more likely biofuels scenario has a continual ramp-up of production. Figure 2 shows the CIRFs for the effect of repeated negative acreage shocks on crop prices. For repeated negative one million corn acreage shocks, corn prices initially rise \$0.04 per bushel, but then continue to rise, peaking at \$0.10 per bushel higher than they would have been, absent any acreage shocks. For repeated negative one million shocks to non-crop farmland, corn prices initially fall \$0.02 per bushel, but then continue to fall. For scenarios 1 and 2 with soybeans, the cumulative effect is also much larger than the one-time shock effect. Two caveats apply to the CIRF results. First, they inevitably have very large standard errors, and the results after the first few periods must be interpreted with caution. Second, CIRFs won't capture changes in expectations. That is, after many periods a repeated "shock" could be incorporated into market expectations, altering the underlying data-generating process and no longer constituting a true "shock."

2.6 Robustness Checks

As described in the modeling section, a number of robustness checks are considered. For instance, a far sparser SVAR (with only five equations) is considered for both crops and both scenarios. Results are quite similar (figure 3). A negative one million acre shock to corn production raises prices by \$0.07. For soybeans, the initial price increase is \$0.23 per bushel for a one million soy acre shock. Decreasing non-crop farmland (holding own acreage constant) lowers corn prices by \$0.06 and soy prices by \$0.11. Next, a log/log specification is estimated for both the main and the sparse models. For scenario 1 with corn, the results are quite similar. For scenario 2, the shape of the IRF is similar but shifted up. For scenario 1 with soy, the results are considerably dampened; however for scenario 2, the results are similar to the those in the linear specification.

The SVAR is also estimated with additional lags allowed in the system (figure 3). For scenario 1 with corn, the estimated IRF is quite similar for two lags. With three lags the initial effect is similar but the effect in later periods is unstable (and implausible). For scenario 1 with soy, the general effect is similar but (implausible) oscillations appear for both two and three lags. For scenario 2, the effect on corn prices is not robust to the inclusion of additional lags. The effect on soy prices is robust to two lags but not three. However the BIC-selected model is one lag for both scenarios and both crops. Moreover there is no theoretical reason to expect additional lags to be relevant.

Next the SVAR is estimated with different diffusion indexes (figure 3). The above results were for the inclusion of the primary diffusion indexes, which loaded mainly onto input prices (supply index) and fairly evenly across US GNP and southern hemisphere agricultural production. The model is also estimated with the secondary diffusion indexes, which loaded primarily onto loan prices (supply index) and Argentine corn and soy production. The results are very similar to those in the main specification (figure 3).

One robustness check uses an additional variable in each specification, to allow for crop rotation. Soy acreage is thus included in the corn specifications, and corn acreage in the soy specifications. Results are nearly identical (figure 3). This is not surprising, since the determinants of soy acreage were already included in the corn equations and the determinants of corn acreage in the soy equations.

As described previously, a few other robustness checks use slightly different data. The main specification is estimated with yields per harvested (rather than planted) acres. Scenario 2 is estimated with a positive shock to total cropland, rather than a negative shock to non-crop farmland. Finally, data on principal crops (rather than total crops) is used in scenario 2, to allow for the inclusion of the year 2007. For all three checks, results are fairly similar to the main specification.

Finally, the SVAR is estimated with a varying time period. The same SVAR is estimated ten times, with 41 years included in each estimation (i.e., 1958 to 1998, 1959 to 1999, etc.). As can be seen from the estimated IRFs (figure 4), the results are quite robust to varying the time frame. For the two corn specifications, the estimation with smaller windows actually shows a larger response. This appears to be because the larger window includes 2007, an anomalous year, which is not included in the smaller time frames. Thus the inclusion of 2007, in which a number of supply and demand shocks hit commodity markets, may bias the results towards zero. Our scenario 1 results, which include 2007, could accordingly be interpreted as a lower bound.

2.7 Conclusion

Using a dynamic system of simultaneous equations, we explore the impacts on crop prices of changes in land use. We develop a structural vector autoregression model, allowing us to analyze impulse response functions and forecast error variance decompositions. These econometric tools, common to macroeconomic applications, provide elegant descriptions of the dynamics of the agricultural production process. We find significant and sustained increases in corn and soybean prices when crop acreage is removed. As described above, this is equivalent to a shock to the residual supply for food. For a reduction in corn area of one million acres, we estimate a corn price increase of \$0.04 per bushel. The last year of our sample (which ends in 2007) saw by far the largest increase in corn acreage dedicated to ethanol - approximately 6 million acres. Real corn prices for this year paid to farmers rose by 88 cents. For 2007, our model therefore predicts that 27% percent of this price increase can be accounted for by the increased area planted to corn for ethanol. Extending our sample further is not currently feasible, yet we can provide a simple calculation for the 2007/08 year. USDA estimates indicate that area planted to corn for ethanol rose by 4.5 million acres and the price paid to farmers rose by \$1.27. Our coefficient estimates based on data up to 2007 would indicate that 14% of this price increase can be attributed to ethanol. If we take a longer run perspective and average the shares over the years 2001-2008, our estimates explain 16% of the year to year fluctuation in prices due to changes in corn acreage planted to ethanol.

For a negative shock in soybean acreage, we find a price increase of \$0.23 per bushel. This much larger magnitude is partly explained by the fact that US soybean acreage is a much larger share of world production than is US corn acreage. We also find significant and sustained decreases in crop prices when own acreage is held constant and total cropland is increased. A 1 million acre increase in US cropland leads to an approximately \$0.06 to \$0.11 decrease in corn and soybean prices. What sets our model apart from most is that the results extend to the production of biofuels besides corn ethanol. Any biofuel feedstock that is grown on land previously dedicated to corn will increase corn prices; this is crucial as the US considers the production of second- and third-generation biofuels.

A number of caveats should be mentioned. First, the magnitudes we see depend on the US share of world production. If this were to change substantially, we might expect a different multiplier. Second, the pathways for the observed responses are only hypothesized. For a scenario in which crop acreage is removed, it is intuitive that the crop's price will rise. Supply is constrained by the removal of acreage, a crucial input in the production process, and demand for the food crop has not changed. The economic rationale behind the second scenario is as follows: redirecting production of biofuels from food crops to second-generation crops will shift the demand for corn inwards, resulting in a drop in corn prices.

Our findings open up a number of possible extensions and future research projects. Scientists and policymakers have expressed hope that new biofuels feedstocks will be grown on land that does not compete with food crops, thus avoiding the effects of biofuels on food prices. Our evidence is suggestive that, if these additional crop lands released pressure from the corn market, corn prices could indeed decline. Verifying this, by analyzing the causal pathways at work, will be crucial as biofuels policies move forward.

2.8 Tables for Chapter 2

Table 1: Unit Root Tests

	Augmented Dickey-Fuller Test with Time Trend		**KPSS Test with Time Trend**
	test statistic	p-value	test statistic
Corn futures	-2.953	0.146	0.179
Soy futures	-2.663	0.252	0.353
Corn price	-2.752	0.215	0.171
Soy price	-2.588	0.285	0.320
Corn acreage	-4.090	0.007	0.101
Soy acreage	-1.149	0.920	0.448
Farmland	0.345	0.996	0.613
Noncrop farmland	-1.485	0.834	0.458
Corn yield	-8.601	0.000	0.076
Soy yield	-6.972	0.000	0.184
Supply Diffusion Index 1	-4.428	0.002	0.115
Supply Diffusion Index 2	-6.973	0.000	0.061
Demand Diffusion Index 1	-8.754	0.000	0.090
Demand Diffusion Index 2	-8.800	0.000	0.045

Notes: All tests use one lag. The null hypothesis of the ADF test is that the variable contains a unit root. The null hypothesis of the KPSS test is that the variable is trend stationary. A Phillips-Perron test with time trend gives similar results, as do ADF and PP tests without time trends. Exceptions are yields, where the null is (not surprisingly) not rejected in the tests without trends. The 5% critical value for the null hypothesis in the KPSS test with time trend is 0.146.

Table 2.1: Forecast Error Variance Decompositions (One-Step Ahead)
Scenario 1, corn

		impulse variable								
		corn futures	soy futures	supply index	corn yields	farmland	corn acr.	demand index	corn harvest price	soy harvest price
response variable	corn futures	1	0	0	0	0	0	0	0	0
	soy futures	0	1	0	0	0	0	0	0	0
	supply index	0.007	0.008	0.985	0	0	0	0	0	0
	corn yields	0.000	0.019	0.002	0.979	0	0	0	0	0
	farmland	0.084	0.046	0.190	0	0.680	0	0	0	0
	corn acreage	0.006	0.046	0.025	0	0	0.923	0	0	0
	demand index	0.032	0.029	0.105	0.001	0.027	0.004	0.80	0	0
	corn harvest price	0.058	0.040	0.164	0.058	0.087	0.068	0.015	0.509	0
	soy harvest price	0.002	0.022	0.236	0.035	0.071	0.025	0.015	0	0.594

Notes: Forecast error variance decomposition tells us the percentage of the forecasting error for a variable due to a specific shock at a given horizon. For instance, the estimate of 0.068 in the corn acreage column, corn harvest price row, tells us that shocks to corn acreage explain 6.8% of the one-step ahead forecast error in corn harvest prices. Note that, since this FEVD is structural, the shocks are orthogonalized.

Table 2.2: Forecast Error Variance Decompositions (One-Step Ahead)
Scenario 2, corn

		impulse variable								
		corn futures	soy futures	supply index	corn yields	corn acr.	non-crop farmland	demand index	corn harvest price	soy harvest price
response variable	corn futures	1	0	0	0	0	0	0	0	0
	soy futures	0	1	0	0	0	0	0	0	0
	supply index	6E-05	0.011	0.989	0	0	0	0	0	0
	corn yields	0.002	0.045	0.005	0.947	0	0	0	0	0
	corn acreage	0.001	0.001	0.051	0	0.947	0	0	0	0
	non-crop farmland	0.088	0.000	0.013	0	0	0.898	0	0	0
	demand index	0.029	0.013	0.107	0.001	0.024	0.028	0.797	0	0
	corn harvest price	0.058	0.017	0.208	0.055	0.009	0.040	0.032	0.581	0
	soy harvest price	0.001	0.021	0.221	0.051	0.021	0.145	0.023	0	0.52

Notes: Forecast error variance decomposition tells us the percentage of the forecasting error for a variable due to a specific shock at a given horizon. For instance, the estimate of 0.040 in the non-crop farmland column, corn harvest price row, tells us that shocks to non-crop farmland explain 4% of the one-step ahead forecast error in corn harvest prices. Note that, since this FEVD is structural, the shocks are orthogonalized.

Table 2.3: Forecast Error Variance Decompositions (One-Step Ahead)
Scenario 1, soy

		impulse variable								
		corn futures	soy futures	supply index	soy yields	farmland	soy acr.	demand index	corn harvest price	soy harvest price
response variable	corn futures	1	0	0	0	0	0	0	0	0
	soy futures	0	1	0	0	0	0	0	0	0
	supply index	0.003	0.005	0.992	0	0	0	0	0	0
	soy yields	0.018	0.001	0.001	0.980	0	0	0	0	0
	farmland	0.100	0.024	0.192	0	0.685	0	0	0	0
	soy acreage	0.022	0.252	0.004	0	0	0.722	0	0	0
	demand index	0.017	0.049	0.136	0.004	0.026	0.033	0.736	0	0
	corn harvest price	0.116	0.053	0.178	0.036	0.002	7E-05	0.022	0.593	0
	soy harvest price	0.001	0.040	0.241	0.056	0.035	0.086	0.014	0	0.528

Notes: Forecast error variance decomposition tells us the percentage of the forecasting error for a variable due to a specific shock at a given horizon. For instance, the estimate of 0.086 in the soy acreage column, soy harvest price row, tells us that shocks to soy acreage explain 8.6% of the one-step ahead forecast error in soy harvest prices. Note that, since this FEVD is structural, the shocks are orthogonalized.

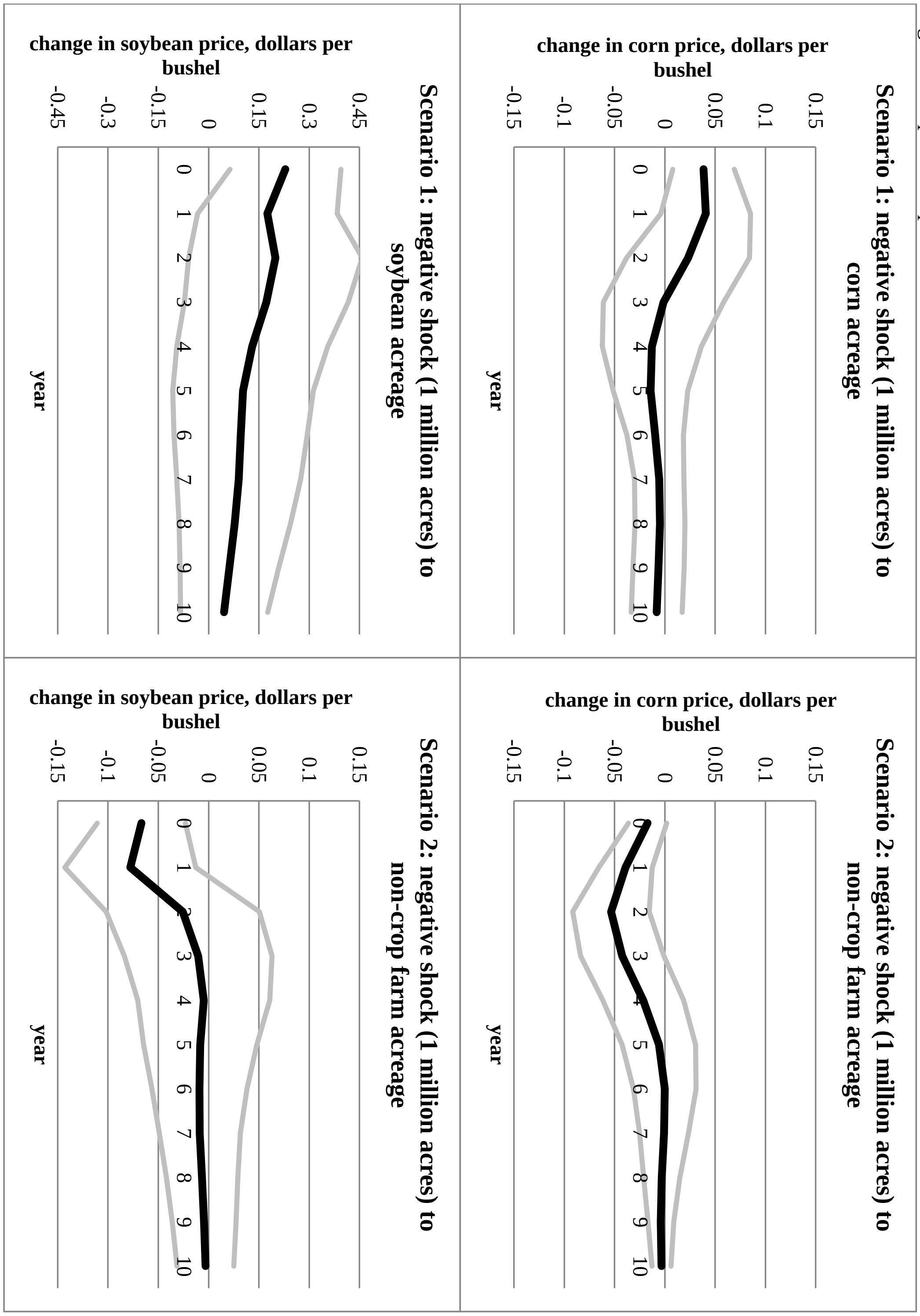
Table 2.4: Forecast Error Variance Decompositions (One-Step Ahead)
Scenario 2, soy

		impulse variable								
		corn futures	soy futures	supply index	soy yields	soy acr.	non-crop farmland	demand index	corn harvest price	soy harvest price
response variable	corn futures	1	0	0	0	0	0	0	0	0
	soy futures	0	1	0	0	0	0	0	0	0
	supply index	0.005	0.005	0.991	0	0	0	0	0	0
	corn yields	0.003	0.004	0.007	0.987	0	0	0	0	0
	corn acreage	0.015	0.086	2E-04	0	0.898	0	0	0	0
	non-crop farmland	0.109	3E-05	1E-05	0	0	0.891	0	0	0
	demand index	0.006	0.003	0.100	0.000	0.053	0.009	0.828	0	0
	corn harvest price	0.061	0.016	0.183	0.038	0.119	0.240	0.040	0.304	0
	soy harvest price	0.003	0.027	0.255	0.075	0.001	0.107	0.049	0	0.484

Notes: Forecast error variance decomposition tells us the percentage of the forecasting error for a variable due to a specific shock at a given horizon. For instance, the estimate of 0.107 in the non-crop farmland column, soy harvest price row, tells us that shocks to non-crop farmland explain 10.7% of the one-step ahead forecast error in soy harvest prices. Note that, since this FEVD is structural, the shocks are orthogonalized.

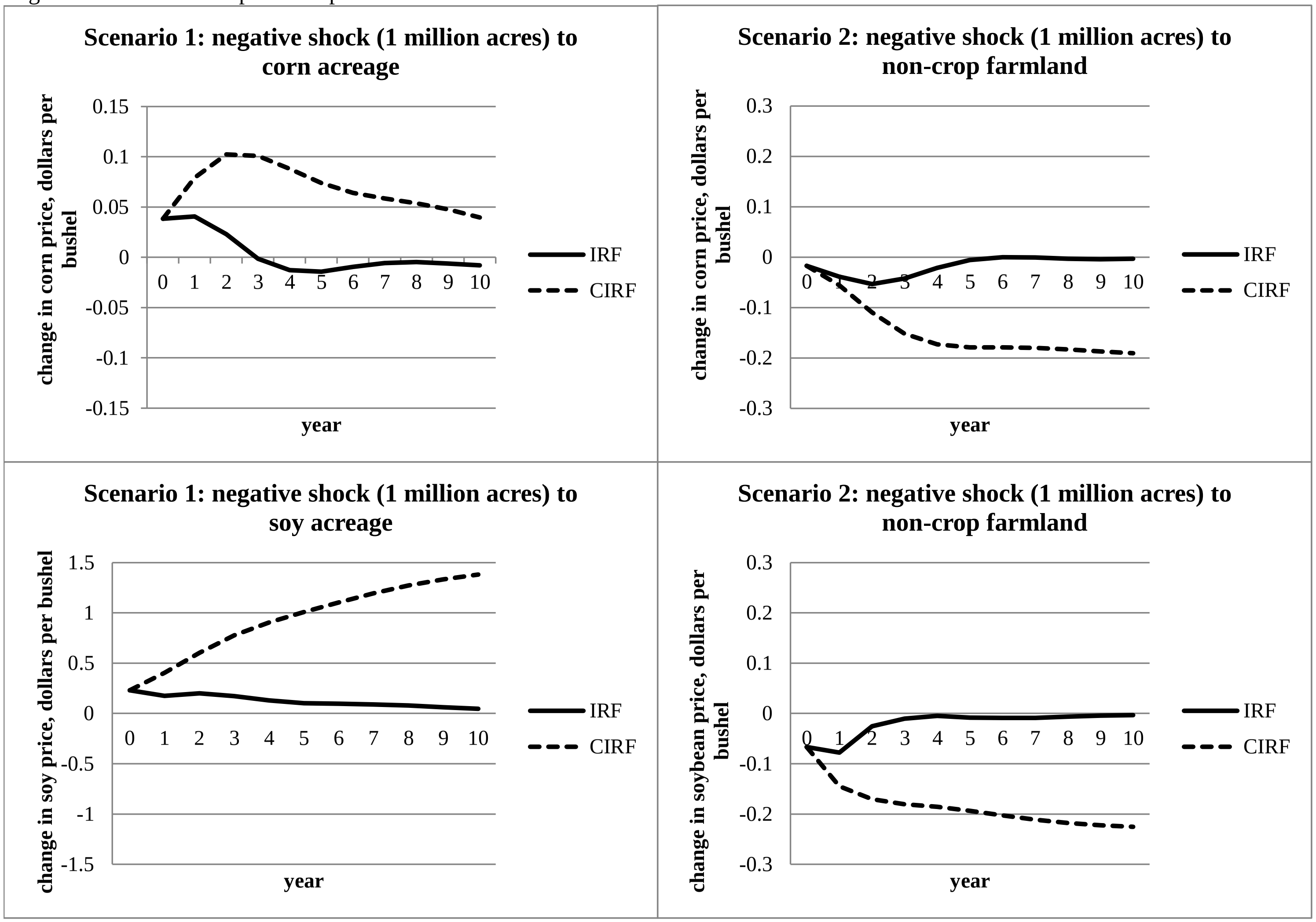
2.9 Figures for Chapter 2

Figure 1: Impulse Response Functions



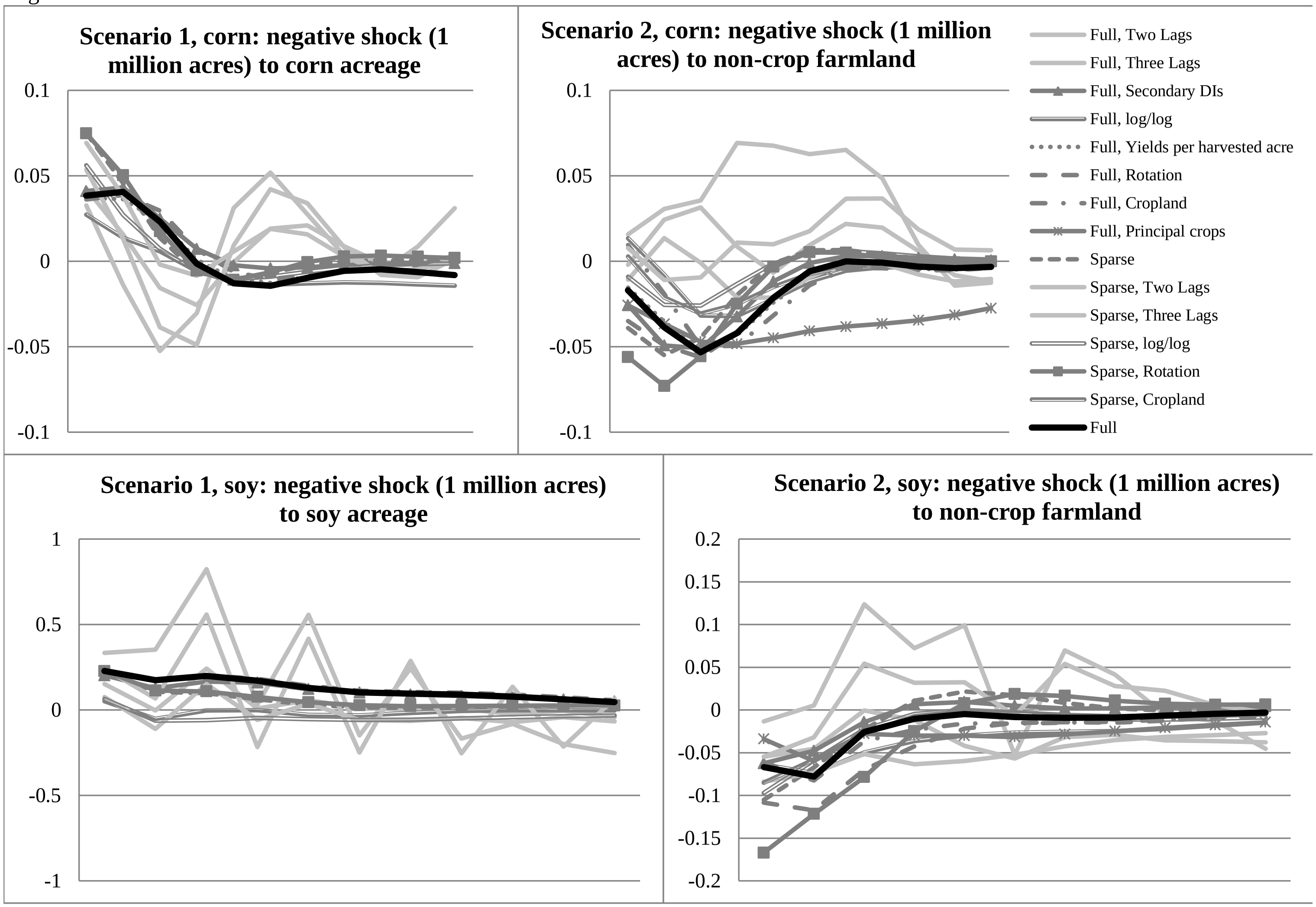
Notes: Impulse response functions trace out the effect of exogenous shocks on realizations of the random variables across time. Since the system is structural, the shocks are orthogonalized. The y-axis shows the change in the U.S. corn or soybean price, and the x-axis shows the period following the shock (which occurs in period zero). Confidence intervals are at the 95% level.

Figure 2: Cumulative Impulse Response Functions



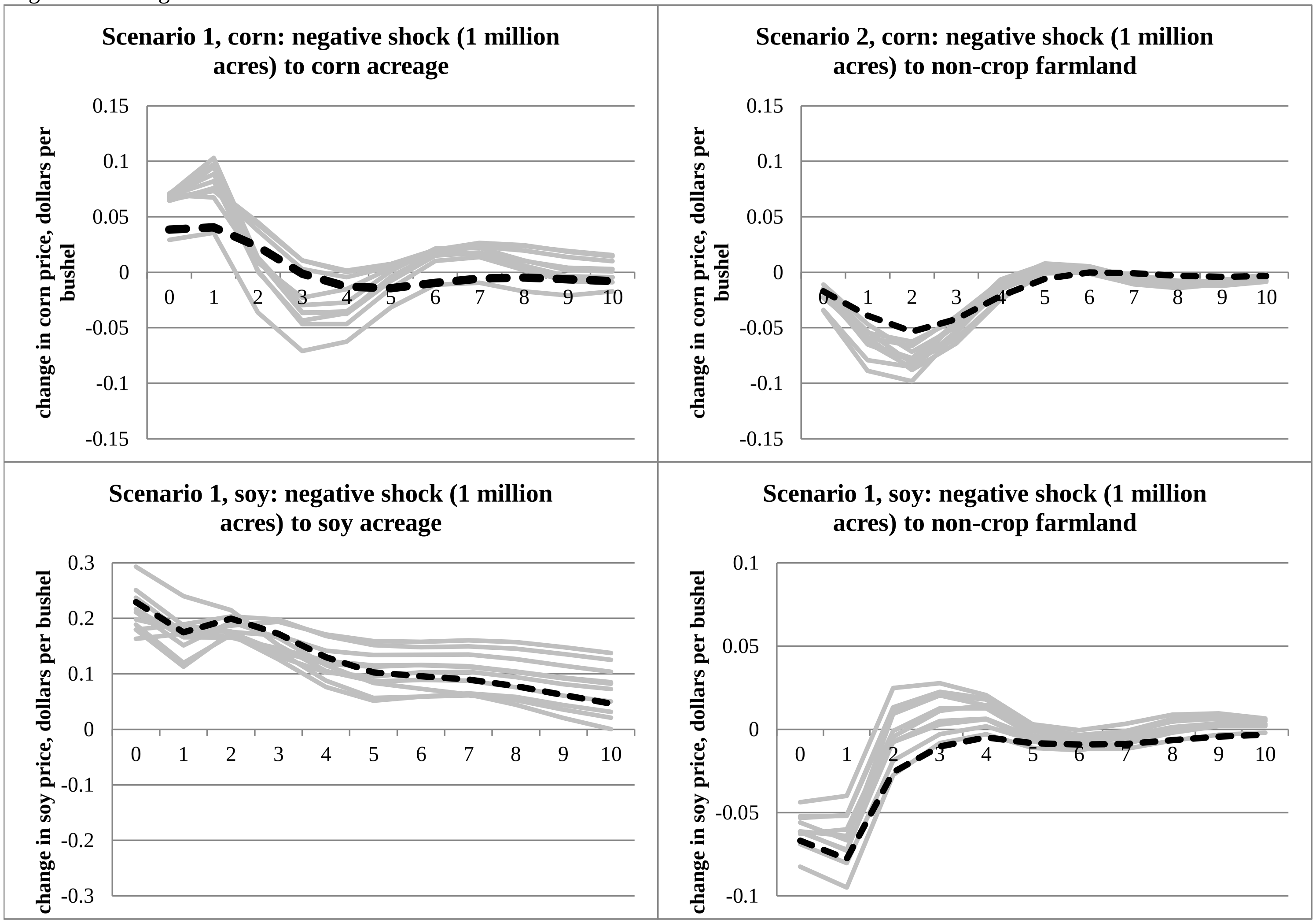
Notes: Impulse response functions trace out the effect of exogenous shocks on realizations of the random variables across time. Since the system is structural, the shocks are orthogonalized. The cumulative IRF shows the path of the crop price over time for repeated acreage shocks. The y-axis shows the change in the U.S. corn or soybean price, and the x-axis shows the period following the first shock (which occurs in period zero).

Figure 3: Robustness Checks



Notes: Impulse response functions trace out the effect of exogenous shocks on realizations of the random variables across time. Since the system is structural, the shocks are orthogonalized. The y-axis shows the change in the U.S. corn or soybean price, and the x-axis shows the period following the shock (which occurs in period zero).

Figure 4: Rolling Time Frame



Notes: The estimated IRF for the entire sample is shown with a dashed line. The rolling window estimation results are shown in gray. The same SVAR is estimated ten times, with 41 years included in each estimation (i.e., 1958 to 1998, 1959 to 1999, etc.). The y-axis shows the change in the U.S. corn or soybean price, and the x-axis shows the period following the shock (which occurs in period zero).

Chapter 3

Biofuels and Land Use Change: Sugarcane and Soybean Acreage Response in Brazil[1]

This chapter is published as: Hausman, Catherine. 2012. "Biofuels and Land Use Change: Sugarcane and Soybean Acreage Response in Brazil." *Environmental and Resource Economics* 51(2): 163-187. The final publication is available at link.springer.com.

3.1 Introduction

With production of biofuels expanding worldwide, concerns about their effect on land use are growing in importance. As the acreage devoted to biofuels crop production expands, it can compete with cropland used for food or with natural ecosystems. Where biofuels production competes with food cropland, it may lead to rising food prices; where it competes with natural ecosystems, it can lead to the loss of biodiversity as well as to the loss of valuable carbon sinks (such as forest or grasslands). Accordingly, economists, scientists, and policymakers are increasingly concerned with understanding the factors that affect trade-offs in the use of land resources. Land use changes are thought to be both direct (land is needed to grow the corn or sugarcane itself) and indirect (dedicating land to these crops pushes up worldwide commodity prices, leading to increased acreage conversion of other crops or in other countries).

The objective of this paper is to investigate the responsiveness of agricultural cropland to changes in crop prices. This paper focuses on Brazil, where agricultural land cover has expanded in recent decades, and where significant trade-offs in land resources are faced: between cropland, pasture, and natural ecosystems such as forest and grasslands. In partic-

[1]I thank Maximilian Auffhammer, Peter Berck, and Joshua Hausman for their invaluable help. I also thank Avery Cohn, Ethan Ligon, Gordon Rausser, Alex Solis, Sofia Villas-Boas, Lunyu Xie, and Carlos Young for excellent comments. This work was carried out under the support of the Energy Biosciences Institute. All opinions are my own and do not represent those of the Energy Biosciences Institute. All errors are mine.

ular, Brazil has seen a large and self-sustaining sugarcane ethanol industry develop at the same time that vast regions of the Amazon and cerrado (a bio-diverse savannah) have been cleared. A host of factors impact land use change in Brazil, so to what extent have sugarcane and soybean prices (both related to biofuels production) been to blame? To date there has been little empirical evidence to support or refute the land use change hypothesis in Brazil, a country with over thirty years of sugarcane ethanol production.

I fill this gap by using a classic agricultural production model of adaptive expectations to estimate how Brazilian acreage changes as international commodity prices change. The parameters of interest are directly estimated using a comprehensive dataset and panel data econometric techniques. I estimate the crop-price elasticity of acreage for eleven major Brazilian crops, focusing especially on sugarcane (a biofuels feedstock) and soybeans. The elasticities can also be estimated across six regions, including areas traditionally devoted to sugarcane and/or soybeans as well as areas of ecological interest. Particularly important are the Amazon rainforest, of which over 15% has been deforested (Pfaff et al. 2007), the cerrado, of which approximately 50% has been cleared (Klink and Machado 2005), and the Mata Atlántica (Atlantic forest), of which over 90% has been cleared.

Regional estimates are important in several ways. The elasticity estimates for sugarcane in ecologically important regions such as the Amazon and cerrado can inform environmental impact assessments of biofuels. The estimates for sugarcane in other regions may provide information about how land use in the Amazon and cerrado could react to changing conditions over time (for instance, additional infrastructure in the Amazon). The estimates for soybeans serve several important functions. First, they provide a benchmark against which to compare sugarcane-propelled land use changes; such benchmarks are critical for life cycle analysis. Second, soybeans have been considered as another potential biofuels feedstock. Finally, empirical evidence strongly suggests that corn ethanol production in the US raises international soybean prices (Searchinger et al. 2008). As such, the impact of land use changes stemming from US corn ethanol production could include Brazilian soybean expansion.

Using county-level data from 1973 to 2005, I find that the short-run price elasticity of Brazilian sugarcane acreage is close to zero, whereas the estimates for the elasticity of soybean acreage are 0.26 for spot prices and 0.63 for futures prices. Regional acreage responses for sugarcane vary little. The highest regional acreage responses for soybeans are in the Center-West and the cerrado, areas that border the Amazon. These findings highlight the differences in land use changes across crops and across regions, and hopefully will inform the ways we analyze the impact of biofuel production.

The most important contribution of this paper is the examination of the links between crop prices and land use changes. Also, I conclude that regional differences matter a great deal in land conversion analysis, which should be taken into account in life cycle assessments and conservation policy-making. I also control for selection of a specific crop in a given region via a Heckman estimator. The factors affecting this selection are of interest in their own right, since in Brazil the agricultural frontier is still expanding. While countries like the United States face an agricultural land constraint, Brazil still has vast areas that could be converted to crop production. This potential conversion has implications for environmental and land-use policies.

Finally, I compare illustrative projections for my estimation with results from other papers, including Searchinger et al. (2008) and Nelson and Robertson (2008). The former uses a Computable General Equilibrium (CGE) model to estimate world-wide land use change from biofuels, and the latter uses a simulation based on satellite data. I also compare my results to those obtained by Nagavarapu (2010), who uses a general equilibrium model that incorporates labor decisions, sugarcane ethanol production decisions, and the import/export market. The advantage of that approach is that, as it integrates labor, agricultural, and non-agricultural markets, it allows for policy simulation. The necessary limitation, then, is that is unable to allow for dynamic agricultural production decisions or disaggregation across fine geographic units. My work complements these previous papers by incorporating detailed data on agricultural land use and crop prices, and by directly estimating the agricultural parameters of interest in the Brazilian context.[2]

3.2 Background

Sugarcane and soybeans are chosen because together they represent almost 50% of all planted acreage in Brazil, and because their production has grown remarkably over the last three decades. Other important crops are beans, cocoa, coffee, corn, cotton, manioc, oranges, rice, and wheat (figure 1).

Brazil has a wide array of ecosystems, and agricultural production varies immensely from region to region. Factors such as degree of mechanization, input costs, soil quality, and climate all impact the crops chosen in each region. Historically, sugarcane production has been centered along a narrow strip of Northeastern coastal land and in the rich agricultural lands of the Southeast (particularly São Paulo state) and the South. Production in these regions is characterized very differently: "Sao Paulo stands on its own technologically as a production region. It is by far the lowest-cost region with the most mechanized technology and least dependence upon labour. The two states in the north use significantly higher levels of labour input and utilize intermediate amounts of machinery and chemicals when compared to the central and southern states. These relative intensities are mirrored in the price elasticity [of inputs] results across each region..." (Rask 1995). Anecdotal evidence suggests that Rask's conclusions still hold.

Ethanol production from sugarcane is centered in the Southeast. The Brazilian government launched the PROALCOOL program in 1975 in an attempt to reduce dependence on foreign oil. It was also hoped that the program would stabilize sugar production. As sugarcane ethanol has grown more competitive with gasoline, government support has been reduced. Currently both anhydrous and hydrous ethanol are produced; the former is blended with gasoline. Approximately half of sugarcane produced in Brazil is refined into ethanol (Balcome and Rapsomanikis 2008).

Soy production is centered in the South and Southeast, but is increasingly pushing north and west into the Amazon (table 1). Given the regional production differences, we can expect

[2]Other research using global CGE models include Gurgel et al. (2007) and Melillo et al. (2009).

varying price elasticities of land in different areas. Hence I analyze six regions of interest separately; states are grouped as Southeast, South, Center-West, Amazon-Border, Amazon-Interior, and Coastal Northeast.[3] The Southeast is defined for this paper as the states of São Paulo, Minas Gerais, Rio de Janeiro, and Espírito Santo. This region has rich farmland and a long history of sugarcane production, although corn and soy are also currently important (table 1). The Southeast is the primary ethanol producing region. The South, another area traditionally devoted to agriculture, is defined as Paraná, Santa Catarina, and Rio Grande do Sul. The Center-West is defined as the states of Mato Grosso, Mato Grosso do Sul, and Goiás (also including the Distrito Federal); this is a region with heavy and rapidly increasing soybean production. Amazon-Border is defined as the states of Mato Grosso, Tocantins, and Maranhão. The legal Amazon as defined by the Brazilian government includes these three states (in addition to six others). These states form a large part of the cerrado, a bio-diverse savannah that has experienced rapid land clearing (Klink and Machado 2005). The most important crops aside from soybeans are rice and corn. Amazon-Interior is defined as the states of Rondônia, Acre, Amazonas, Roraima, Pará, and Amapá. These are also included in the legal Amazon, as defined by the Brazilian government, but have seen less deforestation. Transportation costs are extremely high in this area, limiting agricultural expansion. Other limiting factors include soil quality and topography (Fearnside 2001). Crop acreage is still quite small, but rice, manioc, and corn are more important than other crops. Finally, the Coastal Northeast is defined as Sergipe, Alagoas, Pernambuco, Paraíba, Rio Grande do Norte, and Ceará. The coastal areas of these states have traditionally been devoted to sugarcane. In this region, there may be little room to expand total agricultural production. Sugarcane and soybean acreage has grown, in contrast to other important crops such as beans, corn, and cotton. Clearly production is markedly different across regions, and there may be precision to be gained by analyzing each region separately.

3.3 Data

My research uses county-level data[4] (totaling 3,659 units) from 1973 to 2005 obtained from the government agency IPEA (Instituto de Pesquisa Econômica Aplicada) on acreage, price, and yield of sugarcane, soybeans, beans, cocoa, coffee, corn, cotton, manioc, oranges,

[3]Brazil is traditionally divided into five regions, which are slightly different from the divisions used in this paper. The traditional division groups together states with different land types, ecosystems, and production histories. While estimates could be found for these traditionally defined regions, they would presumably be noisier than those found with the division that I use.

[4]The data are collected by the Instituto Brasileiro de Geografia e Estatística (IBGE 2002), and published by IPEA, the Instituto de Pesquisa Econômica Aplicada (IPEA). Since counties changed borders repeatedly during this time frame, I use AMC (minimal comparable areas) aggregated data from IPEA. Counties whose borders did not change are each considered one unit, and counties whose borders changed are aggregated to minimal comparable areas. For instance, if an area considered county x in time t is later divided into counties x_1 and x_2, the minimal comparable area is x for all time periods. If a portion of county x is moved into county y, the areas of the two counties are combined into one minimal comparable area z for all time periods. Data are used for the whole country, representing 3,659 AMCs.

rice, and wheat.[5] The IPEA price data are available yearly and use harvest-time prices paid to producers. The data collection agency for the above variables, the governmental agency IBGE (Instituto Brasileiro de Geografia e Estatística), provides detailed explanations of data collection and quality control (IBGE 2002). Prices are deflated by the GDP deflator. Additionally, futures data are gathered on sugarcane, soybeans, cocoa, coffee, corn, cotton, and wheat. The futures prices used are planting time prices for futures maturing at harvest time. Crop seasons vary by region; accordingly I use region-specific prices (see the crop calendar in appendix A) and obtain some cross-sectional variation. Unfortunately, crop futures have only recently begun to be traded on the Brazilian market, so US-traded futures are used. I convert these to real Brazilian currency units using the nominal Dollar-Real exchange rate and the Brazilian GDP deflator. Appendix B presents summary statistics. These data allow a detailed, panel data approach to estimating region-specific crop elasticities in Brazil.

3.4 Model

There is a large literature on acreage response, beginning with Nerlove (1956). Nerlove's pioneering work on agricultural supply assumed that farmers make land use decisions according to their expectations of crop prices and input prices. Under the assumption that expectations are adaptive, farmers formulate beliefs about prices as a function of observed prices in previous periods. This assumption leads to a partial-adjustment model, in which acreage in the current period is a function of last year's acreage and of prices during the previous harvest. Nerlove's model has since been applied widely, with many theoretical and empirical adjustments. For instance, futures prices have been included as an additional input into the price expectation function (Gardner 1976). An additional explanation has been developed for the inclusion of the lagged dependent variable as an explanatory variable: it reflects producer inertia, arising from costs of adjustment for switching crops. Risk variables have also been introduced (e.g. Lin and Dismukes 2006). Finally, econometric innovations have allowed more recent work to use panel data, whereas early work relied on time-series data. For very complete reviews of the theoretical and empirical acreage response literature, see Askari and Cummings (1977) and Nerlove and Bessler (2001).

Hence I use a partial-adjustment framework, with current acreage a function of last year's acreage and prices at last year's harvest. Following much of the literature, I include price risk and yield risk. To allow for crop rotation, I also include the lagged acreage of other crops. Acreage, price, yield and risk variables are logged. This leads to the following specification:

[5] Price and yield data are available only for AMCs that harvested a given crop, so for all counties with missing price or yield data, state-level averages are used. Where state-level averages are unavailable, national averages are used. This will not lead to observations being dropped, as there is still year-to-year variation. This is preferable to simply dropping the observations, which would lead to a selection bias (as all observations dropped will be from counties with zero production). Dropped outliers are also replaced with state-level averages (fewer than 20 observations out of 120747). Finally, planted acreage is unavailable, so I use harvested acreage.

$$A_{i,t} = \gamma A_{i,t-1} + \beta_1 PC_{i,t-1} + \mathbf{PS_{i,t-1}}\beta_\mathbf{2} + \mathbf{AS_{i,t-1}}\beta_\mathbf{3} + \beta_4 YLD_{i,t-1} + \beta_5 PrR_{i,t-1} + \beta_6 YR_{i,t-1} + \lambda t_s + \upsilon_i + \mu_t + \varepsilon_{i,t} \quad (3.1)$$

where A is logged county-level crop (sugarcane or soybean) acreage planted in hectares. PC is the log of county-level price of the crop.[6] **PS** is a vector of logged county-level prices of substitute crops (beans, cocoa, coffee, corn, cotton, manioc, oranges, and rice). **AS** is a vector of logged county-level substitute crop acreages.[7] PrR is the log of county-level price risk. Following Chavas & Holt (1990) and Lin & Dismuskes (2006), price risk is defined as the weighted sum of the squared deviation of average price from current price for the previous three periods: $PR_t = 0.5(AP - P_{t-1})^2 + 0.3(AP - P_{t-2})^2 + 0.2(AP - P_{t-3})^2$ where average price is calculated as a moving average: $AP = \frac{P_{t-1}+P_{t-2}+P_{t-3}}{3}$. YLD is the log of county-level crop yield. YR is the log of county-level yield risk: the weighted sum of residuals from the expected yields regression, using the same weights as are used for price risk. Linear trends are represented by t and vary by state s; I check the robustness of the results to the inclusion of quadratic or cubic trends as well as to the exclusion of any state-level trends. County-level fixed effects are captured by υ, and year fixed effects (such as national policy changes) are captured by μ. It is assumed that $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2)$ is a random disturbance and the acreage series is stationary (as seen in the results section, unit root tests find evidence for $|\gamma| < 1$). Standard errors are clustered by state to allow for spatial and temporal correlation.[8]

Some work has suggested that futures prices may more accurately reflect producer expectations than do lagged farmgate prices. Accordingly a similar model is used to estimate price effects in which producers form decisions based on both local spot prices (reflecting the basis) and international futures prices.[9] US-traded futures (at the Chicago Board of Trade

[6] I also tried quadratic and cubic specifications of crop prices. The overall elasticity for the coefficient on own spot price in the sugarcane equation is similar, as are the coefficients on both own spot and own futures prices in the soy equation. The coefficient on own futures price in the sugarcane equation becomes more negative. The coefficients on the higher order terms are generally not significant, with the exception of the quadratic term on own spot price, indicative of possible negative returns to scale, for both sugarcane and soy. Results are available upon request.

[7] It has been suggested that I include other acreage variables. For instance, the amount of pasture land may be of interest. Unfortunately, this is not measured on an annual basis. Head of cattle are measured, but the density of cattle on land is endogenous and varies widely. One could also include the total amount of land in a county, but this would be captured by the fixed effects. Finally, one could include the amount of cropland not captured in the vector of substitute crop acreages (i.e., total cropland minus the sum of the croplands included in **AS**. I do this for equation (3.2). The coefficient on the "other acreage" variable is positive and significant in the soy estimation (perhaps pointing to the usefulness of agricultural infrastructure) but small and not significant in the sugarcane estimation, and it does not change the magnitude or sign of the other reported coefficients.

[8] Despite the presence of prices (which may contain unit roots) on the right hand side, the panel data set eliminates the need for standard errors that are complicated functionals of Wiener processes (Banerjee 1999). Note also that a stationary acreage series is quite intuitive; there is a natural, physical limit to the amount of acreage that can be devoted to agriculture, and as such the variable cannot be explosive.

[9] It has been suggested that I use the basis (equal to the local spot price minus the futures price) and the futures price, rather than the spot price and futures price. The basis can be negative, and as such the

and IntercontinentalExchange) are used, as Brazilian crop futures were not traded for much of this time period.[10] This leads to the following specification:

$$\begin{aligned} A_{i,t} = \gamma A_{i,t-1} + \beta_1 PC_{i,t-1} + \mathbf{PS_{i,t-1}}\beta_2 + \mathbf{AS_{i,t-1}}\beta_3 + \beta_4 YLD_{i,t-1} + \beta_5 PrR_{i,t-1} \\ +\beta_6 YR_{i,t-1} + \beta_7 FPC_{r,t-1} + \mathbf{FPS_{r,t-1}}\beta_8 + \beta_9 FPrR_{r,t-1} + \lambda t_s + v_i + \varepsilon_{i,t} \end{aligned} \tag{3.2}$$

where FPC is the log futures price at planting time for maturity at harvest time of the crop of interest, and **FPS** is a vector of logged futures price of substitute crops (both are subscripted r because they vary by region rather than county). $FPrR$ is the log of futures price risk, calculated similarly to the spot price risk. While some cross-sectional variation is present in futures prices (because of the differing planting and harvesting times across regions), it is small enough to warrant dropping year effects. State-level linear trends are left in the specification to allow for some exogenous trends; I again check the robustness of the results to quadratic and cubic trends and to the exclusion of all trends.

OLS with fixed effects (OLS-FE) is biased in the presence of lagged dependent variables. While consistent, it is asymptotically valid only as the number of time periods (T) grows large (Nickell 1981), rather than as the number of individuals grows large. Intuitively, the bias arises because, after the within transformation has been applied, the lagged dependent variable is correlated with the error term.[11] The bias diminishes only as T grows large because of the incidental parameters problem. That is, as the number of individuals grows larger, the number of parameters also increases, because of the individual fixed effects.

A number of solutions have been proposed, mainly involving Generalized Method of Moments estimators. GMM models have several drawbacks; while asymptotically efficient, they can produce large standard errors. GMM methods for dynamic panel date typically use lagged values of the dependent variables as instruments. Accordingly, the number of moment conditions increases rapidly with T. Also, the use of lagged values of the variables as instruments is problematic (because of the weak instruments problem) when the coefficient on the lag of the dependent variable is close to unity (Stock et al. 2002), as we might expect it to be in an acreage specification.

An alternative to GMM is a bias-corrected fixed effects estimator (OLS-FE-K), such as that proposed by Kiviet (1995). The intuition behind this method is that one can directly estimate the size of the finite-sample bias by beginning with a preliminary GMM estimator.

---

log/log specification used would be inappropriate. Alternatively, one could use the ratio of the spot to the futures price. This is numerically equivalent to the specification used: the coefficient on this ratio would be equal to β_1, and the coefficient on the futures price would be equal to $\beta_1 + \beta_7$ in equation (3.2).

[10] One might also suspect that oil prices would affect the production decision. However, oil prices are correlated with a host of macroeconomic variables, and thus their effect cannot credibly be identified. Since oil prices are not correlated with sugar futures prices in this sample (the estimated correlation coefficient is -0.08), their exclusion should not bias the coefficients on the sugar futures price. While soy futures prices and oil futures prices are correlated in this sample, a robustness check in which oil futures prices were included did not substantially change the coefficients on soy spot or futures prices.

[11] For instance, if the original equation to be estimated is $y_{i,t} = \gamma y + \alpha_i + \varepsilon_{i,t}$, then the within transformation becomes $y_{i,t} - (\frac{1}{T}\sum_{t-1}^{T} y_{i,t}) = \gamma[y_{i,t-1} - (\frac{1}{T}\sum_{t=1}^{T} y_{i,t})] + \varepsilon_{i,t} - (\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t})$. Since $(\frac{1}{T}\sum_{t=1}^{T} y_{i,t})$ is correlated with $(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t})$, the transformed lagged dependent variable is correlated with the transformed error term.

This can then be subtracted from the OLS-FE coefficients. This method can be thought of as a middle ground between some bias with low variance (OLS-FE) and unbiasedness with high variance (GMM). Which of the three methods (OLS with fixed effects, GMM, or bias-corrected OLS with fixed effects) is optimal depends on the particular sample and data-generating process to be studied. Attanasio et al. (2000) find that for a sufficiently large T, OLS bias is small, and thus preferred to a GMM estimator. Monte Carlo studies have found that for moderately-sized T (e.g., approximately 30), Kiviet's bias-corrected estimate is preferred (Judson and Owen 1999, Kiviet 1995).

Accordingly, I report results for both OLS with fixed effects and Kiviet's bias-corrected estimator.[12] For the latter, I use the consistent Anderson-Hsiao instrumental variables estimator (1982) to approximate γ and σ_ε^2 in equations (3.1) and (3.2). This estimator uses $y_{i,t-2}$ as an instrument in the first differences model with no intercept; the instrument is correlated with the transformed dependent variable $y_{i,t} - y_{i,t-1}$ but uncorrelated with the transformed error term $\varepsilon_{i,t} - \varepsilon_{i,t-1}$. The estimates of$\gamma$ and σ_ε^2 are then used to estimate the bias in the OLS-FE model (the approximation of which is derived by Kiviet), and the approximated bias is subsequently subtracted from the OLS-FE coefficients. The standard errors are bootstrapped.

Next, note that the long-run price elasticity of acreage can also be estimated from equations (3.1) and (3.2). Iterative substitution gives us that the long-run price elasticity of acreage is equal to $\beta_1^{LR} = \frac{\beta_1}{1-\gamma}$.[13] The standard errors are approximated using the delta method.

We may expect different elasticities for the intensive and extensive margins. That is, it may be easier or harder to bring new land into agricultural production than to switch from one crop to another. Unfortunately data are not available on land use prior to agricultural use. That is, I cannot observe whether acres devoted to sugarcane were previously devoted to another crop, a non-agricultural use, or were undeveloped. What can be observed is whether or not a given county plants a crop within a given year. Hence a two-step procedure is used. The first step is a participation model in which producers choose whether or not to farm a given crop. In the second step, producers choose the amount of acreage to devote to that crop, conditional on having chosen to produce it. Unfortunately the decisions of individual farmers are unobservable. However, what can be estimated is the marginal effects by county. This leads to the following specification for the first step:

$$\begin{gathered} D_{i,t} = f(D_{i,t-1}, PC_{i,t-1}, \mathbf{PS_{i,t-1}}, YLD_{i,t-1}, PrR_{i,t-1}, YR_{i,t-1}, FPC_{r,t-1}, \mathbf{FPS_{r,t-1}}, \\ FPrR_{t,t-1}) \end{gathered} \tag{3.3}$$

where D is a dummy variable equal to one if the crop of interest is planted in year t and equal to zero otherwise. The model uses a logit specification and allows for fixed effects at the county level. Standard errors are clustered by county. The second step is of the same

[12] I have also tried GMM methods. They do not perform as well as OLS or Kiviet's corrector (as expected), but results are available upon request.

[13] Note that a static between estimator for the long-run elasticity, of the sort proposed by Pirotte (1999), is inappropriate because key assumptions, such as homogeneous effects, are not met.

form as equation (3.2), but is estimated only for counties that plant the crop of interest in period t and accordingly conditions on the inverse Mills ratio (the ratio of the probability density function to the cumulative density function) to control for selection.

Finally, indirect land use changes merit analysis. The above specifications have a clear basis in the producer's decision making, but they do not measure the indirect land use changes that may arise from changes in crop prices. That is, a policy-maker may be more concerned with changes in total crop acreage than with changes in acreage devoted to a specific crop. The cross-price elasticities given in specifications (1) through (3) hint at this, but do not give a complete picture. In a very simple world, with only a few crops, the above specifications would accurately reflect all agricultural land use changes. Given the large number of crops grown in Brazil, however, it may be desirable to analyze a model of total crop acreage. The following model is one way of doing this:

$$AT_{i,t} = \gamma AT_{i,t-1} + \mathbf{PC_{i,t-1}\Phi_1} + \mathbf{FPC_{r,t-1}\Phi_2} + \mathbf{YLD_{i,t-1}\Phi_3} + \mathbf{PrR_{i,t-1}\Phi_4} + \mathbf{YR_{i,t-1}\Phi_5} + \mathbf{FPrR_{r,t-1}\Phi_6} + \lambda t_s + v_i + \varepsilon_{i,t} \tag{3.4}$$

where AT is total planted acreage (aggregating across all crops). $\mathbf{PC_{i,t-1}}$ is a vector of all local spot crop prices (varying by county i) and $\mathbf{FPC_{s,t-1}}$ is a vector of crop futures prices (varying by region r). $\mathbf{YLD_{i,t-1}}$ is a vector of expected yields for all crops. $\mathbf{PrR_{i,t-1}}$, $\mathbf{YR_{i,t-1}}$, and $\mathbf{FPrR_{r,t-1}}$ are vectors of price risk, yield risk, and futures price risk respectively, for all crops. Expected yields, spot and futures price risk, and yield risk are calculated in the same way as for specifications (1) and (2).

3.5 Results and Analysis

For specification (1) estimates are presented in table 2. The sugarcane price elasticity of acreage is estimated to be approximately zero. The soybean price elasticity of acreage is estimated to be 0.24 with a standard error of 0.04. This conforms with the hypothesis that sugarcane acreage is stickier than soybean acreage. When sugarcane prices change, farmers do not respond in the short-run by changing their sugarcane acreage. However, if soybean prices double, soybean acreage increases by 24%. Specification (1), when adjusted with Kiviet's bias corrector, gives very similar estimates. Presumably the bias is negligible because of the moderately large T. Other specifications also show a negligible bias, and accordingly only OLS is reported for all other specifications.[14] For specification (2), which includes crop futures prices, results also suggest that sugarcane acreage is slower to respond to price incentives than is soybean acreage. The coefficients on spot prices and futures prices in the sugarcane model are again 0. For the soybean model, the coefficient on spot prices rises slightly to 0.26 (with a standard error of 0.06) and the coefficient on the soybean futures prices is 0.63 (with a standard error of 0.24). The standard errors on futures prices are likely

[14]One exception is the coefficient on own yield in the sugarcane equations. As can be seen in table 2, the coefficient dropped from 0.40 in the OLS specification to 0.27 in the Kiviet specification. A similar drop in this coefficient was seen for equation (3.2) and for the regional results.

higher (in this specification and the following specifications) because of the limited cross-sectional variation in futures prices and because spot and futures prices are generally highly correlated. One may wish to consider a linear combination of the two coefficients, since one would expect spot and futures prices to move together. This coefficient is 0.89 for soybeans, with a standard error of 0.29. If both own spot and own futures prices doubled, sugarcane acreage would essentially not increase, whereas soybean acreage would increase 89%.

The auto-regressive estimates for the two crops are significantly different. The coefficient on soybeans is 0.6 (with a standard error of 0.02) while the coefficient on sugarcane is 0.75 (with a standard error of 0.01). The auto-correlation coefficient could be indicative of inertia; a coefficient closer to unity indicates that it will take longer to reach equilibrium after a shock. In these national estimates, sugarcane acreage is “stickier” than soybean acreage, both in the response to shocks in prices and in the auto-correlation estimate. However, this result should be treated with caution, as the autocorrelation coefficient might also reflect unobservable dynamic factors.[15] The long-run elasticities also show a dramatic difference between sugarcane and soybeans: the price elasticity of sugarcane is again zero, whereas the elasticity for soybeans is 0.6 to 0.7 for the spot price and 1.6 for the futures price. Again, these should be interpreted with caution, as they are dependent on the estimated autocorrelation coefficient.

Own yield displays a positive coefficient for sugarcane (0.27 to 0.40) with a standard error of 0.05. Own yield also has a positive coefficient (0.34 to 0.37) for soybeans, with a standard error of 0.1. Thus farmers appear to be leveraging the possibility of higher profits by increasing acreage when they expect yields to go up.

The results for state-level trends are presented in figure 2. For sugarcane, the largest value is 0.08, for the states of Amazonas and Roraima (both are Amazon-Interior); other values above 0.05 are Mato Grosso (Center-West), Tocantins (Amazon-Border), and Santa Catarina (South). The largest negative values are -0.04 for Maranhao (Amazon-Border) and -0.03 for Goias (Center-West). All of the above are significant at the 1% level. For soybeans, the state-level trends show a wider range and are generally higher. The largest values are 0.36 for Rondonia (Amazon-Interior), 0.20 for Roraima (Amazon-Interior), 0.17 for Tocantins (Amazon-Border), 0.15 for Mato Grosso (Center-West), and 0.13 for Goias (Center-West). The only negative values are for Santa Catarina (South) and Rio Grande do Sul (South). All of the above are significant at the 1% level. Thus, quite aside from price responses, soybean acreage is generally growing faster than that of sugarcane. Both crops have the highest secular trends in regions of ecological importance, such as the Amazon and the Center-West.

As mentioned in the modeling section, I examine the robustness of the results to alternative trend specifications. The full results are not presented, in the interest of space, but are available upon request. First I estimate equation (3.2) with quadratic and cubic trends. In the sugarcane equation, all estimated coefficients of interest (those reported in table 2) remain quite similar whether a quadratic trend or cubic trend is used. Some of the coefficients change in the soybean equation, but their sign does not. For soybeans with a quadratic trend, own spot price goes from 0.26 to 0.25, and the own futures price goes from

[15] I am grateful to an anonymous reviewer for this point.

0.63 to 0.77. For soybeans with a cubic trend, own spot price goes from 0.26 to 0.18 and own futures goes from 0.63 to 0.67. Thus the combined spot+futures elasticity remains close to the previously estimated 0.89. I also estimate equation (3.2) with no trends or year effects. The sugarcane price estimates are still not significantly different from zero, and the yield and lagged acreage coefficients are similar. In the soybean equation, the spot price, futures price, and yield estimates rise somewhat (spot price from 0.26 to 0.3; futures from 0.63 to 0.77; yield from 0.34 to 0.43).

Various specification tests are used to determine whether the above form of the model is correct. Unit root tests find evidence for $|\gamma| < 1$. A Hausman test supports the use of fixed effects rather than random effects, as does an F-test on the intercept terms. A BIC selection criterion indicates that one lag is optimal.[16]

Next I consider the effects of price shocks on the participation response versus the conditional level response. Table 3 shows the results for a logit specification in which the dependent variable is a dummy for county-level participation in sugarcane or soybean production. The results again show that soybean acreage responds more to crop price shocks than does acreage. Columns (3) and (6) gives elasticities calculated at the sample mean values of the independent variables. The elasticity for the local spot price and futures price of sugarcane is negligible. For soybeans the spot price elasticity is 0.22 and the futures price elasticity is 0.44. As described above, futures and spot prices generally move together. Hence the linear combination of the two effects is more relevant than either effect in isolation. The combined effect for sugarcane acreage of spot and futures prices is negligible, whereas the combined effect for soybeans is 0.66. Thus we see that soybeans move into new counties more rapidly in response to price shocks than does sugarcane. Own yields are again positive and statistically significant, with an elasticity of 0.08 for sugarcane and 0.15 for soybeans. These results are robust to several alternative specifications such as a probit specification or excluding futures prices.

Table 4 shows how the level of crop production within a county producing sugarcane or soybeans responds to price shocks. The elasticity for the sugarcane spot price is 0.01 and is not significant, and the elasticity for the sugar futures price is -0.02. The elasticity for the soybean spot price is 0.15 and the elasticity for the futures price is 0.65; both are significant at the 1% level. Thus the combined effect of the two prices is 0.8. This response for soybeans is similar to the unconditional response given in table 2 (0.89) and the participation response in table 3 (0.66). Thus new soybean production occurs both because new counties begin production and because production expands in counties already growing the crop.

A policy-maker may be more concerned with the effect of crop prices on total planted acreage, rather than on acreage devoted to a specific crop. This is a useful starting point for considering indirect land use changes, and estimation results are given in table 5. The highest elasticities are for corn and soybean futures. As the price of soybean futures doubles,

[16] Another variable of potential interest in this model is the effect of government production policies, so I consider state-level government expenditures on agriculture, energy, and transportation. These variables lead to endogeneity, as agricultural producers may be able to influence the level of government expenditures. I try several instrumental variables, but the estimated coefficients are unstable. Generally the own-price estimates on the crops do not change. More detailed data on government expenditures is needed to answer this question, but the price effects do not appear to be sensitive to the omission of this variable.

total crop acreage increases by 27%. The coefficient on corn futures is 0.18, but is not statistically significant. The elasticity is approximately zero for sugar. These results match the conclusion drawn above that soybean price shocks are extremely important to agricultural acreage response; in fact, they are more important than any other crop price analyzed.

This total impact is a function of own-price responses as well as cross-price responses, which are summarized in table 6. Equation (3.2) is estimated for all of Brazil's other major crops (beans, cocoa, coffee, corn, tree cotton, upland cotton, manioc, oranges, rice, and wheat) in addition to sugarcane and soybeans, analyzed above. The strongest own-price responses, aside from soybeans, are for beans (elasticity of 0.3), cotton (0.2 for tree cotton spot price, 0.1 for upland cotton spot price, and 0.9 for upland cotton futures price) and wheat (elasticity of 0.2 for spot price and 0.08 for futures price). However, there has been no evidence to date that these crop prices will be of importance in biofuels production scenarios. The soybean futures price has a significant effect on a number of other crops (elasticities of 1.1, 0.79, 0.66, and 0.21 for beans, coffee, corn, and manioc respectively). Negative effects are also observed for a few crops (elasticities of -0.46 for oranges and -0.13 for wheat). The sugarcane futures prices have economically and statistically significant effects on three crops (elasticities of -0.25 for upland cotton, -0.30 for wheat, and -0.35 for rice). This evidence suggests that changes in soybean prices will have broader impacts on agricultural production than will changes in sugarcane prices.

Interesting spatial heterogeneities become apparent in the regional specifications (tables 7 and 8). The highest elasticities for sugarcane acreage response are found in the Amazon-Border, Amazon-Interior, and Coastal Northeast where spot price elasticities remain low but the elasticity for the own futures price is 0.28, 0.14 and 0.20 respectively. The Amazon-Interior results should be interpreted with caution, however, as the standard error is high and the coefficient is not significant at the 10% level. Soybeans also show the highest acreage response in areas of ecological importance. In the Center-West, the spot price elasticity is 0.99 and the futures price elasticity is 2.9; both are statistically significant. In the Amazon-Border the spot and futures price elasticities are 0.70 and 2.51, respectively, and both are statistically significant. In the Amazon-Interior the spot price elasticity is 0.53 and is significant, and the futures price elasticity is -0.50 and is not statistically significant. The coefficients in the Southeast and South are higher than the national coefficients but lower than the estimates in the Amazon and the Center-West, and the response in the Coastal Northeast is essentially zero.[17] Although the coefficients are not statistically significantly different from one another, these results accord with evidence by other researchers that soybeans are expanding most rapidly in areas of environmental concern. The evidence is magnified in the long-term elasticity. Thus soybeans do appear to be contributing to acreage

[17]The coefficients vary somewhat when cubic state-level time trends are used. In the South, the own futures elasticity for sugarcane goes from -0.25 to -0.13 (not statistically significant in either specification). The own price elasticity for soy generally falls when cubic trends are included: from 0.70 to 0.59 in the Amazon-Border, 0.99 to 0.74 in the Center-West, 0.40 to 0.29 in the Southeast, and 0.26 to 0.06 in the South. The own futures price elasticity for soy also falls: 2.5 to 1.9 in the Amazon-Border, 2.9 to 1.6 in the Center-West, and 1.2 and 1.0 in the South. The coefficient on own yield for soy rises in two regions (0.74 to 0.89 in the Amazon-Interior; 0.55 to 0.65 in the South) and falls in one region (0.55 to 0.42 in the Amazon-Border). Full tables available upon request.

conversion along the edges of the Amazon and to heavy land use change in the ecologically important cerrado of the center of Brazil. It should also be noted that if transportation costs decrease in the interior of the Amazon (for example, as roads improve), we might expect the price elasticity of acreage there to rise to the levels of the Center-West or the Amazon-Border. Land use changes in these regions could be expected to accelerate as agricultural production expands.

Finally, I consider the robustness of the results to varying the time frame. Innovations in agricultural technology (for instance, mechanization or new seed varieties) may have changed the production process over time. I estimate equation (3.2) with a rolling time frame of 15 periods with 16 years each (e.g. 1976 to 1991, 1977 to 1992, etc.). The coefficient on lagged acreage falls somewhat in the sugarcane equation, ranging between 0.6 in the later samples to 0.7 in the earlier samples. This implies a decrease (in absolute value) of the long-run elasticities, as the cumulative effect of a shock will decrease. The own spot price for the sugarcane equation, which is 0.01 and not significant for the entire sample, is generally negative and not significant in the smaller time frames. There are a few periods for which the coefficient on own spot price is negative and significant (the late 1980s and 1990s), but for the same periods the coefficient on the futures price is very large. The coefficient on own yield, which is 0.38 for the entire sample, ranges between 0.2 and 0.7 (with the 95% confidence intervals generally including 0.38).

The rolling window for the soy equation also shows interesting heterogeneity across time. As in the sugarcane equation, the coefficient on lagged acreage falls: from 0.6 for the entire sample to a range of 0.4 to 0.5 for the smaller time frames. The coefficient on own spot price also falls, from 0.26 for the entire sample to a range of 0 to 0.18. Most notably, the estimated coefficient falls as the time frame moves forward, indicating that currently-relevant values are lower than the 0.26 estimated over 33 years. The coefficient on the soy futures price also falls in later years. However, the coefficient on soybean yields rises to a high of 0.4 for the late 1990s and 2000s; it appears that these later years are what drove up the coefficient when estimated for the entire time frame. The rise in the yield coefficient may be related to the introduction of improved soybean varieties. Indeed the most relevant time period is likely the later one, implying an own spot price elasticity of less than 0.1 and an own futures price elasticity of less than 0.5 for soybeans and spot and futures price elasticities indistinguishable from zero for sugarcane. For both crops acreage responds strongly to yield shocks in the latest years.

3.6 Illustrative Calculations

A few illustrative projections can be made from the total acreage specification (equation (3.4) and table 5) to understand the impact of biofuels production. Supposing that the spot and futures prices of sugarcane, soybeans and corn increased 10%, total crop acreage would be expected to increase by 4.8% (calculated as 0.1 multiplied by the sum of these coefficients in table 5). During the final period of the sample, 2005, 62 million hectares were planted in Brazil. Thus the previous price scenario, similar to the average annual increase seen over the last five years, would lead to an increase of 3 million hectares of planted acreage in the near term.

Searchinger et al. (2008) predict a 2.8 million hectare increase, albeit for a different price scenario (in which corn prices increase 40%, soy prices 20% and wheat prices 17%). For this scenario, my projections suggest that total crop acreage would increase by 6 million hectares.[18] Another recent paper (Nelson and Robertson 2008) projects a long-run increase of 165 million hectares for a scenario in which corn prices increase 25% and sugar prices increase 10%. For this scenario, my model projects an immediate increase of 3 million hectares and a long-run increase of 5 million additional hectares.[19] This price scenario, however, is problematic as it ignores changes in soy price (which would presumably follow changes in the corn price).

Direct comparisons to the results obtained by Nagavarapu (2010) are difficult. The simulations presented in that paper include changes to U.S. trade policy (not just agricultural prices), and the land use changes include adjustments in pasture land and privately-held forest land, neither of which are incorporated in my model. However, it is striking to note the differences in regional variation between that work and the results presented here. Nagavarapu finds that the largest changes in sugarcane land use generally come from the region near Sao Paulo, whereas I find the largest changes in regions that border the Amazon. It is hard to know whether this differences arises from Nagavarapu's inclusion of labor market constraints, or from the more disaggregated data and the dynamic model that I use.

3.7 Conclusions

The key finding of this study is that Brazilian soybean acreage grows much faster in response to changes in price than does sugarcane acreage, and the difference is particularly marked in regions of ecological importance. With the elasticities estimated above, as sugarcane prices rise, sugarcane acreage doesn't increase in the short run. The key exception is the Amazon-Border, where the elasticity with respect to the futures price is 0.28. However, as soy price doubles, soy acreage increases 26 percent nationally. As the price of soybean futures doubles, soy acreage increases 63 percent. This increase is higher in regions of ecological importance; as soy spot and futures prices double, soy acreage increases 390 percent in the Center-West and 320 percent along the border of the Amazon. Point estimates for the long-run elasticity of sugarcane are also small, while the estimates for soybeans are approximately 0.65 (spot price) and 1.6 (futures price) nationally. The difference between the crops holds when only the latest 15 years of the sample are used, but the coefficients on own prices in the soybean equation become noticeably smaller. This difference between the crops is robust to considering the effect on county-level participation in crop production using a binary dependent variable. The difference also holds when I estimate the effect on the level of acreage given that the county participates in crop production. These two points relate

[18] Equal to the 62 million multiplied by sum of: 0.4 times the corn coefficients (0.007+0.184), 0.2 times the soy coefficients (0.036+0.272), and 0.17 times the wheat coefficients (0.008-0.249).

[19] The short-run increase of 3 million hectares is calculated from the short-run coefficients on sugar and corn spot and futures prices. The long-run increase is calculated by dividing the short-run coefficients by (one minus the autocorrelation coefficient), as described in the modeling section.

to an important finding: expansion of soybean production into new counties has been as important as increasing production within counties already growing the crop. Additionally, total crop acreage responds more heavily to changes in soybean prices than to changes in sugar prices.

A few caveats naturally apply. First, the importance of the Mata Atlántica (Atlantic forest), of which less than 10% of the original ecosystem remains, should not be understated. The data used above is not sufficiently disaggregated to assess the impact of sugarcane and soybean production on the remaining Atlantic forest, but that is not to say the region is not important in the evaluation of agricultural crops and biofuels. Also, this analysis does not attempt to determine directly how the acreage response factors into life-cycle questions such as whether sugarcane ethanol is more carbon-friendly than other fuels. What this paper does give is an estimate of how land use changes vary across regions. More work on acreage conversion and agriculture is clearly needed; economists and policy-makers alike need more information on indirect land use change and on why regional differences are so large. Future work could take advantage of the availability of satellite data (which Nelson and Robertson (2008) are able to use) or could incorporate constraints (legal, topological, etc.). Also, the estimates above are based on the conditions of the time period and regions studied. Clearly the construction of roads or bio-refineries in ecologically important areas will change the response of agricultural producers. Thus, there may be threshold effects – once a crop is profitable enough, large investments may be made in capital and infrastructure, leading to a large change in land use. However, what can be concluded is that land use change in Brazil over the last three decades is probably much more a product of the reaction of soybean production to market prices than the reaction of sugarcane production to market prices. Agricultural acreage can be expected to expand somewhat in the long-term if biofuels production permanently increases sugar prices. However, the response of sugarcane acreage to price changes has been significantly smaller than that of soybean acreage, in both statistical and economic terms.

3.8 Tables for Chapter 3

Table 1: Regional Variation in Crop Acreage (Acreage in Thousands of Hectares)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Southeast		South		Center-West		Amazon - Border		Amazon - Interior		Coastal Northeast	
Crop	**1973**	**2005**	**1973**	**2005**	**1973**	**2005**	**1973**	**2005**	**1973**	**2005**	**1973**	**2005**
Beans	840	631	1066	655	202	193	92	131	16	155	1259	1153
Cocoa	29	21	0	0	0	1	0	1	8	86	0	0
Coffee	6	1814	1	106	23	29	1	18	1	193	22	13
Corn	3107	2486	3944	3687	557	2088	428	1417	65	474	2750	2054
Cotton	660	166	293	57	221	678	110	470	1	0	2262	45
Manioc	291	139	526	283	81	88	243	244	123	478	720	314
Oranges	389	615	28	50	5	7	2	2	1	18	33	88
Rice	1430	149	995	1199	1269	1009	952	1498	108	457	534	1051
Soy	224	1900	3245	7980	146	10336	0	6343	0	159	2218	3475
Sugarcane	1189	3666	147	453	35	535	37	239	18	15	730	957
Wheat	29	72	1786	2125	24	109	0	747	0	0	0	0
All crops*	8970	12500	12984	18157	2725	15692	1945	10607	452	2296	13034	11384

* Includes crops other than those listed

Totals across regions do not sum to national totals, as the regional divisions used are neither exclusive nor comprehensive.

Source: IPEA

Table 2: National Estimates of Equations (1) and (2) for Sugarcane and Soybeans: Measuring Acreage Response in Brazil

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS-FE †	OLS-FE-K ‡	OLS-FE †	OLS-FE †	OLS-FE-K ‡	OLS-FE †
Region	Brazil					
DEPENDENT VARIABLE	sugarcane acreage			soybean acreage		
own acreage (lagged)	0.747	0.791	0.746	0.600	0.646	0.601
	(0.00983)***	(0.00249)***	(0.00958)***	(0.0230)***	(0.00253)***	(0.0228)***
price of sugarcane	-0.00668	-0.00123	0.00889	0.0463	0.0470	0.0343
	(0.0200)	-0.0205	(0.0243)	(0.0331)	(0.0181)***	(0.0320)
futures price of sugar			-0.0174			-0.221
			(0.0963)			(0.0737)***
price of soybeans	0.118	0.127	0.0862	0.236	0.234	0.261
	(0.0477)**	(0.0458)***	(0.0490)*	(0.0430)***	(0.0437)***	(0.0644)***
futures price of soybeans			-0.978			0.628
			(0.166)***			(0.235)***
expected own yield	0.402	0.266	0.378	0.370	0.366	0.336
	(0.0546)***	(0.0287)***	(0.0614)***	(0.119)***	(0.0385)***	(0.117)***
own price risk	0.0115	0.0138	0.00805	-0.00543	-0.00575	-0.00594
	(0.0116)	(0.00571)**	(0.00968)	(0.0114)	(0.00720)	(0.00768)
own futures price risk			0.0407			0.00348
			(0.0118)***			(0.00827)
own yield risk	0.00590	0.00541	0.00603	0.00498	0.00616	0.00593
	(0.00215)***	(0.00190)***	(0.00242)***	(0.00962)	(0.00550)	(0.00713)
Observations	109770	109770	109770	109770	109770	109770
R-squared (Within)	0.579		0.576	0.411		0.410
Spot prices of other crops	Y	Y	Y	Y	Y	Y
Futures prices of other crops	N	N	Y	N	N	Y
Year effects	Y	Y	N	Y	Y	N
County-level effects	Y	Y	Y	Y	Y	Y
State-level linear trends	Y	Y	Y	Y	Y	Y
spot + futures price of sugar			-0.009			-0.187
			(0.106)			(0.0752)**
spot + futures price of soy			-0.892			0.889
			(0.162)***			(0.290)***
Long-run own-price elasticity	-0.026	-0.006	0.035	0.590	0.661	0.654
	(0.0385)	(0.0119)	(0.0384)	(0.0877)***	(0.0511)***	(0.0996)***
Long-run own futures price elasticity			-0.069			1.574
			(0.0463)			(0.896)*

Notes: All variables are logged, so the above estimates are elasticities. Results are similar in columns (3) and (6) if futures prices are included and local prices excluded. All specifications control for local prices and lagged acreage of rice, oranges, cocoa, coffee, corn, upland cotton, tree cotton, manioc, wheat, and beans. In columns (3) and (6), I also control for futures prices of cocoa, coffee, corn, wheat, and cotton. Constants are not reported because of the fixed-effects specification. † Robust, clustered (by state) standard errors in parentheses. ‡ Bootstrapped standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Results are fairly similar if a quadratic or cubic state-level trend is used. Please see the results section of the text for details.

Table 3: Effects on Participation in Production of the Crop of Interest

	(1)	(2)	(3)
DEPENDENT VARIABLE		**sugarcane**	
independent variable	**coefficient**	**standard error**	**elasticity**
own lagged participation	5.04	0.04	0.62
own spot price	-0.07	0.02	-0.01
own futures price	0.00	0.04	0.00
own yields	0.46	0.03	0.08

	(4)	(5)	(6)
DEPENDENT VARIABLE		**soybeans**	
independent variable	**coefficient**	**standard error**	**elasticity**
own lagged participation	3.37	0.05	0.70
own spot price	0.26	0.02	0.22
own futures price	0.51	0.06	0.44
own yields	0.30	0.02	0.15

Logit specification. Dependent variable is a dummy, equal to zero if cane/soy acreage is zero and equal to one otherwise. All specifications control for local prices of rice, oranges, cocoa, coffee, corn, upland cotton, tree cotton, manioc, wheat, and beans; they also control for futures prices of cocoa, coffee, corn, wheat, and cotton. Elasticities are calculated at sample mean values of the independent variables. Standard errors are robust and clustered (at the county level). Results are similar if a probit specification is used or if futures prices are not included. Results are also similar if the dependent variable is a dummy equal to one if cane/soy acreage is greater than 10 hectares. Regional results available upon request.

Table 4: Sugarcane and Soybean Acreage Response Conditional on Acreage Participation in Brazil, 1976-2005

	(1)	(2)
	Heckman	
Region	Brazil	
DEPENDENT VARIABLE	sugarcane acreage	soybean acreage
own acreage (lagged)	0.143	0.0741
	(0.0182)*	(0.0125)
own spot price	0.00943	0.149
	(0.0291)	(0.0477)***
own futures price	-0.0249	0.646
	(0.0314)	(0.107)***
expected own yield	0.360	0.127
	(0.113)	(0.0810)
own price risk	0.00538	0.0144
	(0.00910)	(0.0161)
own futures price risk	0.00871	0.0186
	(0.00511)*	(0.00720)***
own yield risk	-0.000979	-0.00316
	(0.00259)	(0.00485)
Observations	77187	26893
R-squared (Within)	0.271	0.363
Spot and futures prices of other crops	Y	Y
Year effects	N	N
County-level effects	Y	Y
State-level linear trends	Y	Y
own spot + own futures price	-0.015	0.795
	(0.0421)	(0.117)***

Notes: All variables are logged, so the above estimates are elasticities. Observations are limited to those counties that planted sugarcane or soybeans. All specifications control for local prices and lagged acreage of sugarcane, soybeans, rice, oranges, cocoa, coffee, corn, upland cotton, tree cotton, manioc, wheat, and beans; they also control for futures prices of sugar, soybeans, cocoa, coffee, corn, wheat, and cotton. Results are similar if either spot or futures prices are excluded. Constants are not reported because of the fixed-effects specification. Robust, clustered (by state) standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Regional results available upon request.

Table 5: Total Crop Acreage Response in Brazil, 1976-2005

	(1) OLS-FE
Region	Brazil
DEPENDENT VARIABLE	otal crop acreage
total crop acreage (lagged)	0.676
	(0.0595)***
price of beans	0.0610
	(0.0240)**
price of cocoa	-0.00206
	(0.0127)
price of cocoa future	0.0191
	(0.0201)
price of coffee	-0.0451
	(0.0482)
price of coffee future	0.0384
	(0.0442)
price of corn	0.00730
	(0.00791)
price of corn future	0.184
	(0.146)
price of upland cotton	-0.0361
	(0.0275)
price of cotton future	-0.0769
	(0.0406)*
price of tree cotton	0.0209
	(0.0252)
price of manioc	-0.00555
	(0.0128)
price of oranges	-0.0130
	(0.0226)
price of rice	0.0198
	(0.0428)
price of soybeans	0.0358
	(0.0250)
price of soybean future	0.272
	(0.0556)***
price of sugarcane	-0.0119
	(0.0126)
price of sugar future	-0.00342
	(0.0320)
price of wheat	0.00811
	(0.0136)
price of wheat future	-0.249
	(0.0959)***

Observations	109770
R-squared (Within)	0.485
Year effects	N
County-level effects	Y
State-level trends	Y
Expected yield, price risk, yield risk	Y
Long-run sugarcane price elasticity	-0.0367
	(0.184)
Long-run sugar futures price elasticity	-0.011
	(0.18)
Long-run soybean price	0.110
	(0.18)
Long-run soybean futures price elasticity	0.840
	(0.21)***

Notes: All variables are logged, so the above estimates are elasticities. Results are similar if futures prices are included and local prices excluded, or if futures prices are excluded and local prices included. Constants are not reported because of the fixed-effects specification. Robust, clustered (by state) standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%.

Table 6: cross price elasticities

DEPENDENT VARIABLE: acreage of:	beans	cocoa	coffee	corn	upland cotton
own spot price	0.331	0.00282	-0.113	-0.166	0.135
	(0.0765)***	(0.0140)	(0.0594)*	(0.110)	(0.138)
own futures price		-0.0181	0.138	1.119	0.864
		(0.00691)***	(0.115)	(0.596)*	(0.229)***
price of soybeans	0.174	0.00679	-0.0998	0.158	-0.0243
	(0.130)	(0.00897)	(0.140)	(0.0868)*	(0.145)
price of soybean	1.110	-0.0530	0.786	0.663	0.334
future	(0.362)***	(0.0509)	(0.334)**	(0.270)**	(0.234)
price of sugarcane	-0.0182	-0.00312	0.0560	0.00721	0.158
	(0.0663)	(0.00287)	(0.0473)	(0.0684)	(0.0803)**
price of sugar future	-0.00729	-0.0279	-0.116	-0.234	-0.245
	(0.109)	(0.0133)**	(0.1000)	(0.194)	(0.126)*
Observations	109770	109770	109770	109770	109770
R-squared	0.242	0.636	0.488	0.100	0.397

DEPENDENT VARIABLE: acreage of:	tree cotton	manioc	oranges	rice	wheat
own spot price	0.213	-0.0489	-0.0805	0.0935	0.217
	(0.0867)**	(0.0629)	(0.0359)**	(0.116)	(0.0952)**
own futures price					0.0795
					(0.0583)
price of soybeans	0.0198	0.00653	0.0454	0.0894	-0.0819
	(0.0325)	(0.0755)	(0.0921)	(0.0957)	(0.0576)
price of soybean	0.233	0.207	-0.457	0.410	-0.129
future	(0.190)	(0.118)*	(0.226)**	(0.252)	(0.124)
price of sugarcane	0.0274	0.0439	0.0730	-0.00186	0.00750
	(0.0189)	(0.0458)	(0.0478)	(0.0373)	(0.0289)
price of sugar future	-0.0195	-0.0403	-0.143	-0.350	-0.296
	(0.0529)	(0.0717)	(0.101)	(0.112)***	(0.118)**
Observations	109770	109770	109770	109770	109770
R-squared	0.777	0.441	0.584	0.422	0.277

Notes: All variables are logged, so the above estimates are elasticities. All specifications control for own and other lagged acreage, own price risk, own expected yield, own yield risk, county-level effects, and state-level trends. Constants are not reported because of the fixed-effects specification. Robust, clustered (by state) standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Full results available upon request.

Table 7: Sugarcane Acreage Response in Brazilian Regions, 1976-2005

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS-FE					
Region	**Southeast**	**South**	**Center-West**	**Amazon - Border**	**Amazon - Interior**	**Coastal Northeast**
DEPENDENT VARIABLE	sugarcane acreage					
sugarcane acreage (lagged)	0.755	0.716	0.779	0.716	0.731	0.730
	(0.0150)***	(0.0123)***	(0.0386)***	(0.0764)***	(0.0252)***	(0.0171)***
price of sugarcane	0.0365	-0.171	-0.00811	0.0601	0.00531	-0.0381
	(0.0209)*	(0.102)*	(0.0534)	(0.0226)***	(0.154)	(0.0796)
futures price of sugar	-0.0964	-0.248	0.0398	0.277	0.143	0.199
	(0.120)	(0.503)	(0.0867)	(0.0298)***	(0.469)	(0.107)*
expected own yield	0.389	0.720	0.184	0.235	-0.0774	0.393
	(0.0898)***	(0.194)***	(0.208)	(0.381)	(0.307)	(0.0951)***
own price risk	0.0129	-0.0103	-0.0244	0.0137	-0.0236	0.0223
	(0.0139)	(0.0329)	(0.0124)**	(0.0328)	(0.0467)	(0.0301)
own futures price risk	0.0773	0.0886	0.0559	0.0255	-0.118	0.0377
	(0.00992)***	(0.0195)***	(0.0342)	(0.0300)	(0.0350)***	(0.0138)***
own yield risk	0.00549	0.0111	(0.0156	-0.00347	0.0303	-0.00284
	(0.00441)	(0.00776)	(0.0100)	(0.00368)	(0.0123)***	(0.00463)
Observations	42030	17820	6690	5100	3270	23340
R-squared (Within)	0.583	0.54	0.649	0.591	0.578	0.577
Spot and futures prices of other crops	Y	Y	Y	Y	Y	Y
Year effects	N	N	N	N	N	N
County-level effects	Y	Y	Y	Y	Y	Y
State-level trends	Y	Y	Y	Y	Y	Y
spot + futures price of sugar	-0.0599	-0.419	0.03169	0.3371	0.14831	0.1609
	(0.137)	(0.604)	(0.0833)	(0.0433)***	(0.412)	(0.0891)*
Long-run own-price elasticity	0.149	-0.602	-0.0367	0.212	0.0197	-0.141
	(0.0585)**	(0.229)***	(0.172)	(0.286)	(0.0989)	(0.0770)*
Long-run own futures price elasticity	-0.393	-0.873	0.180	0.975	0.532	0.737
	(0.242)	(1.559)	(0.105)*	(0.315)***	(0.836)	(0.263)***

Notes: All variables are logged, so the above estimates are elasticities. Results are similar if futures prices are included and local prices excluded. All specifications control for local prices and lagged acreage of rice, oranges, cocoa, coffee, corn, upland cotton, tree cotton, manioc, wheat, and beans; I also control for futures prices of cocoa, coffee, corn, wheat, and cotton. Constants are not reported because of the fixed-effects specification. Robust, clustered (by state) standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%.

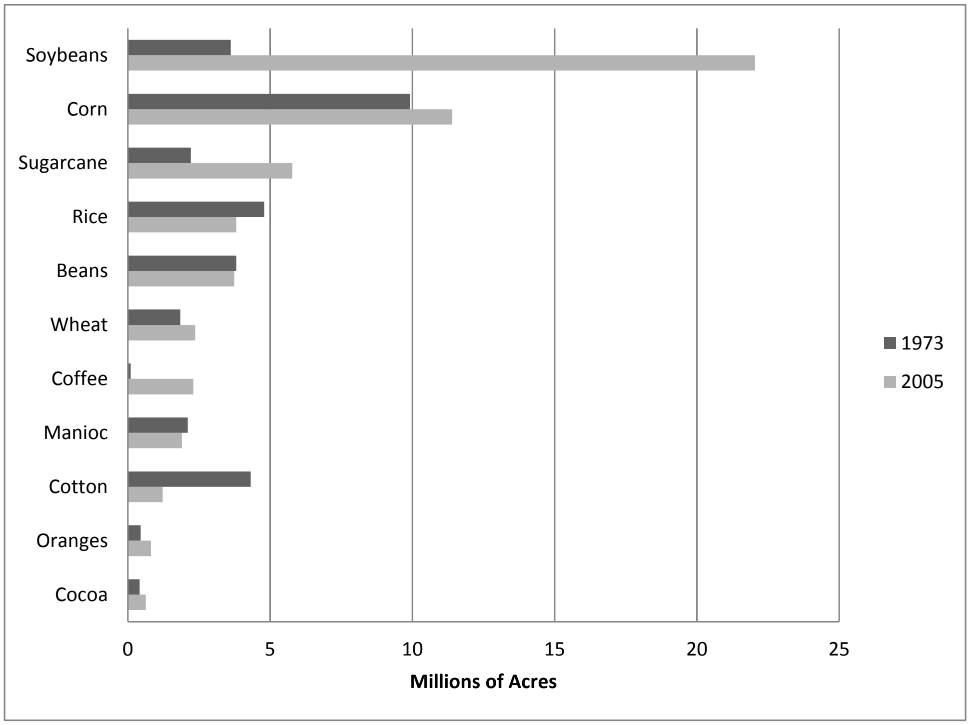
Table 8: Soybean Acreage Response in Brazilian Regions, 1976-2005

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS-FE					
Region	**Southeast**	**South**	**Center-West**	**Amazon - Border**	**Amazon - Interior**	**Coastal Northeast**
DEPENDENT VARIABLE	soybean acreage					
soybean acreage (lagged)	0.564	0.638	0.573	0.654	0.853	0.521
	(0.0285)***	(0.0488)***	(0.0288)***	(0.0784)***	(0.0247)***	(0.0453)***
price of soybeans	0.397	0.260	0.991	0.703	0.528	-0.0169
	(0.219)*	(0.182)	(0.294)***	(0.319)**	(0.151)***	(0.0125)
futures price of soybeans	1.062	1.249	2.930	2.506	-0.495	-0.0782
	(0.349)***	(0.320)***	(0.386)***	(0.730)***	(0.977)	(0.0599)
expected own yield	0.472	0.548	1.570	0.547	0.741	0.0264
	(0.219)**	(0.164)***	(0.432)***	(0.200)***	(0.111)***	(0.0474)
own price risk	0.0211	-0.0438	-0.0373	0.00595	0.0450	0.00209
	(0.0114)*	(0.0272)	(0.0515)	(0.0376)	(0.0510)	(0.00208)
futures price risk	0.0141	-0.0107	-0.0400	-0.0675	0.00867	-0.00684
	(0.00138)***	(0.0379)	(0.0486)	(0.0901)	(0.0353)	(0.00492)
own yield risk	0.00819	-0.0176	-0.0696	-0.0514	0.0495	-0.00301
	(0.00494)*	(0.00591)***	(0.0384)*	(0.0868)	(0.0127)***	(0.0104)
Observations	42030	17820	6690	5100	3270	23340
R-squared (Within)	0.334	0.497	0.475	0.556	0.678	0.255
Spot and futures prices of other crops	Y	Y	Y	Y	Y	Y
Year effects	N	N	N	N	N	N
County-level effects	Y	Y	Y	Y	Y	Y
State-level trends	Y	Y	Y	Y	Y	Y
spot + futures price of soybeans	1.459	1.509	3.921	3.209	0.033	-0.0951
	(0.564)***	(0.337)***	(0.305)***	(1.047)***	(0.898)	(0.0689)
Long-run own-price elasticity	0.911	0.718	2.321	2.032	3.592	-0.0353
	(0.398)**	(0.333)**	(1.535)	(1.650)	(3.604)	(0.0948)
Long-run own futures price elasticity	2.436	3.450	6.862	7.243	-3.367	-0.163
	(1.882)	(2.935)	(6.258)	(15.057)	(22.366)	(0.109)

Notes: All variables are logged, so the above estimates are elasticities. Results are similar if futures prices are included and local prices excluded. All specifications control for local prices and lagged acreage of rice, oranges, cocoa, coffee, corn, upland cotton, tree cotton, manioc, wheat, and beans; I also control for futures prices of cocoa, coffee, corn, wheat, and cotton. Constants are not reported because of the fixed-effects specification. Robust, clustered (by state) standard errors in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%.

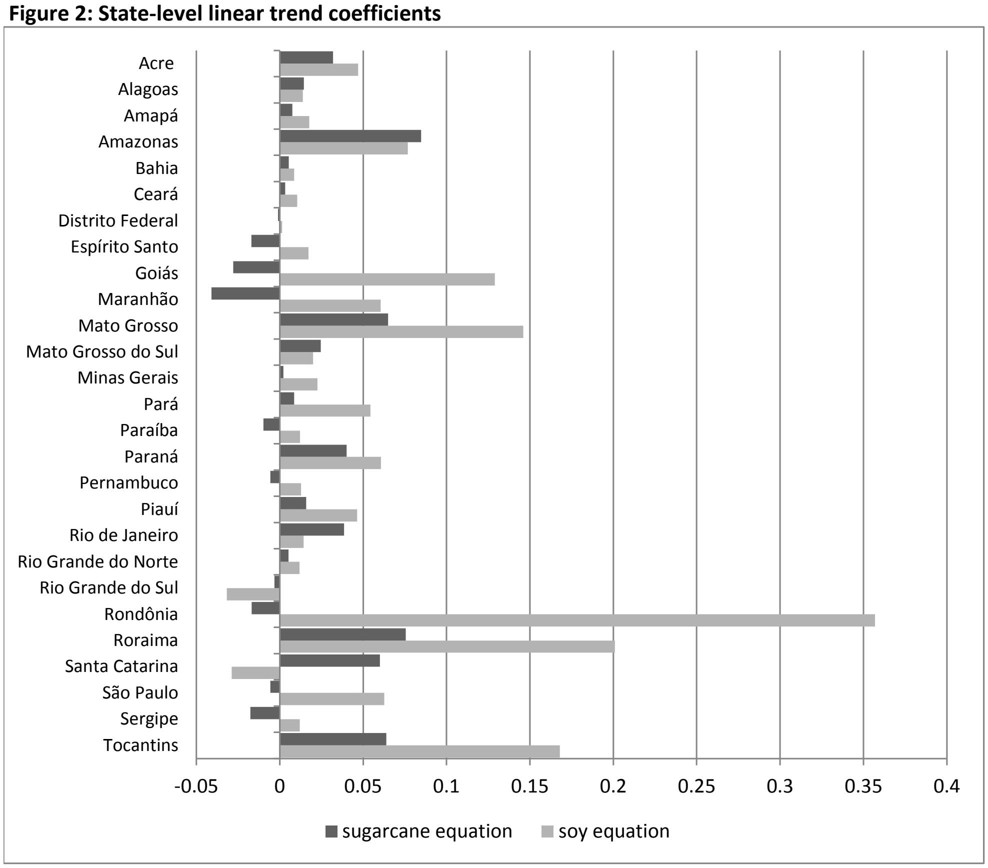
3.9 Figures for Chapter 3

Figure 1: Major Crops of Brazil



Source: IPEA

Figure 2: State-level linear trend coefficients



Full results from these two regressions are presented in columns (3) and (6) of table 2. The regressions estimate the acreage response of sugarcane and soybeans in Brazil and are represented by equation (2) in the text.

Appendix A: Crop Calendars

Center and South Crop Calendar

Crop	Jan		Feb		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
Beans*	H	H	H																P	P	P			H
Coffee																								
Corn*				H	H	H	H	H										P	P	P	P			
Cotton	P	P								H	H	H	H	H										P
Oranges																								
Rice					H	H	H	H													P	P	P	
Soy				H	H	H	H													P	P	P	P	
Sugarcane																								

P denotes peak planting time; H denotes peak harvest time.

*First crop - there may be two or three plantings in a year

Source: Companhia Nacional de Abastecimento, USDA

North and Northeast Crop Calendar

Crop	Jan		Feb		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
Beans*				H	H	H	H															P	P	P
Cocoa**																								
Corn*	P	P							H	H	H	H	H										P	P
Cotton		P	P	P	P							H	H	H	H									
Rice	P	P				H	H	H	H	H	H											P	P	P
Soy	P					H	H	H	H													P	P	P
Sugarcane																								

P denotes peak planting time; H denotes peak harvest time.

*First crop - there may be two or three plantings in a year

**Principal harvest - another cocoa harvest runs from May to September

Source: Companhia Nacional de Abastecimento, USDA

Appendix B: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Entire sample (n = 120747)		Year = 1976 (n = 3659)		Year = 2005 (n = 3659)		
Variable	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Units**
cane acreage	1.08	4.36	0.61	2.77	1.58	5.75	thousands of hectares
soy acreage	2.94	28.51	1.75	14.37	6.02	63.91	
total crop acreage	13.87	49.61	15.32	39.38	16.75	95.96	
county landmass	232.26	1419.88	232.26	1420.07	232.26	1420.07	
participation in sugarcane production	0.71	0.45	0.74	0.44	0.68	0.47	equal to 1 if county planted sugarcane, 0 otherwise
participation in soy production	0.24	0.43	0.23	0.42	0.32	0.47	equal to 1 if county planted soybeans, 0 otherwise
farmgate price of beans	1.31	0.76	2.19	1.00	0.76	0.17	Thousand Reais (Reais in year 2000) per metric ton
farmgate price of cocoa	3.51	2.20	5.35	1.42	2.03	0.15	
farmgate price of coffee	2.47	1.73	4.02	1.09	1.96	0.46	
farmgate price of corn	0.33	0.18	0.43	0.11	0.20	0.05	
farmgate price of tree cotton	1.31	0.65	2.24	0.95	0.65	0.06	
farmgate price of upland cotton	1.19	1.98	1.78	0.57	0.68	0.21	
farmgate price of manioc	0.19	0.14	0.22	0.07	0.14	0.10	
farmgate price of oranges	0.40	0.35	0.47	0.19	0.20	0.09	
farmgate price of rice	0.53	0.30	0.70	0.13	0.30	0.10	
farmgate price of soybeans	0.48	0.26	0.58	0.08	0.30	0.07	
farmgate price of sugarcane	0.04	0.04	0.04	0.01	0.03	0.02	
farmgate price of wheat	0.57	0.33	0.77	0.04	0.22	0.05	
soybean yield	1.86	0.52	1.57	0.37	2.35	0.55	metric tons per hectare
sugarcane yield	46.66	19.65	37.45	15.46	56.07	20.25	

Source: IPEA

Bibliography

[1] Abbott, Philip. C., Christopher Hurt, and Wallace E. Tyner. 2008. "What's Driving Food Prices?" Farm Foundation Issue Report.

[2] Allison, Paul D. and Richard P. Waterman. 2002. "Fixed-Effects Negative Binomial Regression Models." *Sociological Methodology* 32(1): 247-265.

[3] Anderson, T.W. and Cheng Hsiao. 1982. "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics* 18(1): 47-82.

[4] Ando, Amy W., Madhu Khanna, and Farzad Taheripour. 2010. "Market and Social Welfare Effects of the Renewable Fuels Standard." *Handbook of Bioenergy Economics and Policy, Natural Resource Management and Policy* 33(4): 233-250.

[5] Ando, Amy W. and Karen L. Palmer. 1998. "Getting on the Map: The Political Economy of State-Level Electricity Restructuring." RFF Discussion Paper 98-19-REV.

[6] Askari, Hossein and John Thomas Cummings. 1977. "Agricultural Supply Response with the Nerlove Model: A Survey." *International Economic Review* 18(2): 257-292.

[7] Attanasio, Orazio P., Lucio Picci, and Antonello E. Scorcu. 2000. "Saving, Growth, and Investment: A Macroeconomic Analysis Using a Panel of Countries." *Review of Economics and Statistics* 82(2): 182-211.

[8] Averch, Harvey and Leland L. Johnson. 1962. "Behavior of the Firm under Regulatory Constraint." *American Economic Review* 52(5): 1052-1069.

[9] Balcome, Kelvin and George Rapsomanikis. 2008. "Bayesian Estimation and Selection of Nonlinear Vector Error Correction Models: The Case of the Sugar-Ethanol-Oil Nexus in Brazil." *American Journal of Agricultural Economics* 90(3): 658-668.

[10] Banerjee, Anindya. 1999. "Panel Data Unit Roots and Cointegration: An Overview." *Oxford Bulletin of Economics and Statistics, Special Issue* (1999): 607-629.

[11] Banse, Martin, Hans van Meijl, Andrzej Tabeau, and Geert Woltjer. 2008. "Will EU Biofuel Policies Affect Global Agricultural Markets?" *European Review of Agricultural Economics* 35(2): 117-141.

[12] Barnett, Arnold and Mary K. Higgins. 1989. "Airline Safety: The Last Decade." *Management Science* 35(1): 1-21.

[13] Baron, David P. 1989. "Design of Regulatory Mechanisms and Institutions." *Handbook of Industrial Organization* 2(24): 1347-1447.

[14] BBC. 2007. "Will biofuel leave the poor hungry?" http://news.bbc.co.uk/2/hi/business/7026105.stm.

[15] Borenstein, Severin, James B. Bushnell, and Frank A. Wolak. 2002. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review* 92(5): 1376-1405.

[16] Bushnell, James B. and Catherine Wolfram. 2005. "Ownership Change, Incentives and Plant Efficiency: The Divestiture of U.S. Electric Generation Plants." CSEM Working Paper 140.

[17] Bushnell, James B., Erin T. Mansur, and Celeste Saravia. 2009. "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets." *American Economic Review* 98(1): 237-266.

[18] Cha, Kyung Soo and Jeong Hwan Bae. 2011. "Dynamic Impacts of High Oil Prices on the Bioethanol and Feedstock Markets." *Energy Policy* 39(2): 753-760.

[19] Chakravorty, Ujjayant, Marie-Helene Hubert, and Linda Nostbakken. 2009. "Fuel Versus Food." *Annual Review of Resource Economics* 1: 645-663.

[20] Chakravorty, Ujjayant, Marie-Helene Hubert, Michel Moreaux, and Linda Nostbakken. 2011. "Will Biofuel Mandates Raise Food Prices?" University of Alberta, Department of Economics Working Paper No. 2011-01.

[21] Chavas, Jean-Paul, and Matthew T. Holt. 1990. "Acreage Decisions under Risk: The Case of Corn and Soybeans." *American Journal of Agricultural Economics* 72(3): 529-538.

[22] Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1-45.

[23] Commodity Research Bureau. 1956-2003. *Commodity Yearbook*.

[24] Congressional Budget Office. 2009. "The Impact of Ethanol Use on Food Prices and Greenhouse-Gas Emissions." Congress of the United States Congressional Budget Office Pub. No. 3155.

[25] Craig, J. Dean and Savage, Scott J. 2009. "Market Restructuring, Competition and the Efficiency of Electricity Generation: Plant-level Evidence from the United States 1996 to 2006." University of Colorado at Boulder Department of Economics Working

Paper No. 09-06. Accessed from http://dirwww.colorado.edu/Economics/papers/Wps-09/wp09-06/wp09-06.pdf.

[26] David, Paul A., Roland Maude-Griffin, and Geoffrey Rothwell. 1996. "Learning by Accident? Reductions in the Risk of Unplanned Outages in U.S. Nuclear Power Plants After Three Mile Island." *Journal of Risk and Uncertainty* 13(2): 175-198.

[27] Davis, Lucas W. 2012. "Prospects for U.S. Nuclear Power." *Journal of Economic Perspectives* 26(1): 49-66.

[28] Davis, Lucas W. and Catherine Wolfram. 2012. "Deregulation, Consolidation, and Efficiency: Evidence from U.S. Nuclear Power." *American Economic Journal: Applied Economics* 4(4): 194-225.

[29] Eide, Steven A., Dale M. Rasmuson, and Corwin L. Atwood. 2005. "Integrated Initiating Event Performance Indicator." Idaho National Laboratory.

[30] Energy Information Adminstration. 2011. *Electric Power Annual 2009*. Washington, DC: U.S. Department of Energy.

[31] Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram. 2007. "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency." *American Economic Review* 97(4): 1250-1277.

[32] Fearnside, Philip M. 2001. "Soybean Cultivation as a Threat to the Environment in Brazil." *Environmental Conservation* 28(1): 23-38.

[33] Feinstein, Jonathan S. 1989. "The Safety Regulation of U.S. Nuclear Power Plants: Violations, Inspections, and Abnormal Occurrences." *Journal of Political Economy* 97(1): 115-154.

[34] Fowlie, Meredith. 2010. "Emissions Trading, Electricity Restructuring, and Investment in Pollution Abatement." *American Economic Review* 100(3): 837-869.

[35] Galiani, Sebastian, Paul Gertler, and Ernesto Schargrodsky. 2005. "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113(1): 83-120.

[36] Gardner, Bruce L. 1976. "Futures Prices in Supply Analysis." *American Journal of Agricultural Economics* 58(1): 81-84.

[37] Golbe, Devra L. 1986. "Safety and Profits in the Airline Industry." *Journal of Industrial Economics* 34(3): 305-318.

[38] Government Accountability Office. 2008. "Nuclear Safety: NRC's Oversight of Fire Protection at U.S. Commercial Nuclear Reactor Units Could Be Strengthened." Report GAO-08-747. Washington, D.C.: U.S. Government Accountability Office.

[39] Griffin, James M. and Steven L. Puller. 2005. "A Primer on Electricity and the Economics of Deregulation." *Electricity Deregulation: Choices and Challenges*, Griffin and Puller, eds. Chicago: University of Chicago Press.

[40] Gurgel, Angelo, John M. Reilly, and Sergey Paltsev. 2007. "Potential Land Use Implications of a Global Biofuels Industry." *Journal of Agricultural and Food Industrial Organization*. Special Issue, Volume 5: Article 9.

[41] Hanemann, W. Michael, Per-Olov Johansson, Bengt Kristrom, and Leif Mattsson. 1992. "Natural Resource Damages from Chernobyl." *Environmental and Resource Economics* 2(5): 523-525.

[42] Headey, Derek and Shenggen Fan. 2008. "Anatomy of a Crisis: The Causes and Consequences of Surging Food Prices." *Agricultural Economics, Issue Supplement* 39(s1): 375-391.

[43] Hill, Jason, Erik Nelson, David Tilman, Stephen Polasky, and Douglas Tiffany. 2006. "Environmental, Economic, and Energetic Costs and Benefits of Biodiesel and Ethanol Biofuels." *Proceedings of the National Academy of Sciences* 103(30): 11206-11210.

[44] Instituto Brasileiro de Geografia e Estatística (IBGE). 2002. *Pesquisas Agropecuarias, 2nd ed.* Rio de Janeiro.

[45] Instituto de Pesquisa Econômica Aplicada (IPEA), www.ipeadata.gov.br.

[46] Joskow, Paul L. 1997. "Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector." *Journal of Economic Perspectives* 11(3): 119-138.

[47] Joskow, Paul L. and Richard Schmalensee. 1986. "Incentive Regulation for Electric Utilities." *Yale Journal on Regulation* 4(1): 1-49.

[48] Judson, Ruth A. and Ann L. Owen. 1999. "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists." *Economics Letters* 65:9-15.

[49] Kahn, Alfred E. 1988. *The Economics of Regulation: Principles and Institutions*. MIT Press.

[50] Kennet, D. Mark. 1993. "Did Deregulation Affect Aircraft Engine Maintenance? An Empirical Policy Analysis." *RAND Journal of Economics* 24(4): 542-558.

[51] Khanna, Madhu, Amy W. Ando, and Farzad Taheripour. 2008. "Welfare Effects and Unintended Consequences of Ethanol Subsidies." *Review of Agricultural Economics* 30(3): 411-421.

[52] Kiviet, J.F. 1995. "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models." *Journal of Econometrics* 68:53-78.

[53] Klink, Carlos A. and Ricardo B. Machado. 2005. "Conservation of the Brazilian Cerrado." *Conservation Biology* 19(3):707–713.

[54] Laffont, Jean-Jacques and Jean Tirole. 1986. "Using Cost Observation to Regulate Firms." *Journal of Political Economy* 94(3): 614-641.

[55] Lin, Willian and Robert Dismukes. 2006. "Supply Response under Risk: Implications for Counter-Cyclical Payments' Production Impact." *Review of Agricultural Economics* 29(1): 64-86.

[56] Lutkepohl, Helmut. 1993. *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

[57] Mansur, Erin T. 2008. "Measuring Welfare in Restructured Electricity Markets." *Review of Economics and Statistics* 90(2): 369–386.

[58] Massachusetts Institute of Technology (MIT). 2003. "The Future of Nuclear Power: An Interdisciplinary MIT Study." MIT Energy Initiative.

[59] Massachusetts Institute of Technology (MIT). 2009. "Update of the MIT 2003 Future of Nuclear Power: An Interdisciplinary MIT Study." MIT Energy Initiative.

[60] McPhail, Lihong Lu. 2011. "Assessing the Impact of U.S. Ethanol on Fossil Fuel Markets: A Structural VAR Approach." *Energy Economics* 33(6): 1177-1185.

[61] Melillo, Jerry M., John M. Reilly, David W. Kicklighter, Angelo C. Gurgel, Timothy W. Cronin, Sergey Paltsev, Benjamin S. Felzer, Xiaodong Wang, Andrei P. Sokolov, and C. Adam Schlosser. 2009. "Indirect Emissions from Biofuels: How Important?" *Science* 326(5958): 1397-1399.

[62] Mitchell, Donald. 2008. "A Note on Rising Food Prices." World Bank Policy Research Working Paper No. 4682.

[63] Mushtaq, Khalid and P. J. Dawson. 2002. "Acreage Response in Pakistan: A Co-Integration Approach." *Agricultural Economics* 27(2): 111-121.

[64] Nagavarapu, Sriniketh. 2010. "Implications of Unleashing Brazilian Ethanol: Trading Off Renewable Fuel for How Much Forest and Savanna Land?" Working paper.

[65] National Agricultural Statistics Service. www.nass.usda.gov.

[66] National Council on Radiation Protection and Measurements (NCRP). 2009. "Ionizing Radiation Exposure of the Population of the United States." NCRP Report No 160.

[67] Naylor, Rosamond L., Adam J. Liska, Marshall B. Burke, Walter P. Falcon, Joanne C. Gaskell, Scott D. Rozell, and Kenneth G. Cassman. 2007. "The Ripple Effect: Biofuels, Food Security, and the Environment." *Environment* 49(9): 30-43.

[68] Nelson, Gerald C. and Richard D. Robertson. 2008. "Green Gold or Green Wash: Environmental Consequences of Biofuels in the Developing World." *Review of Agricultural Economics* 30(3): 517-529.

[69] Nerlove, Marc. 1956. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities." *Journal of Farm Economics* 38(2): 496-509.

[70] Nerlove, Marc and David A. Bessler. 2001. "Expectations, Information and Dynamics." *Handbook of Agricultural Economics* 1:155-206.

[71] Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49(6): 1417-1426.

[72] Nuclear Regulatory Commission. 2010. "Information Digest." NUREG-1350, Volume 22.

[73] Pfaff, Alexander, Juan Robalino, Robert Walker, Steven Aldrich, Marcellus Caldas, Eustaquio Reis, Stephen Perz, Claudio Bohrer, Eugenio Arima, William Laurance, and Kathryn Kirby. 2007. "Road Investments, Spatial Spillovers, and Deforestation in the Brazilian Amazon." *Journal of Regional Science* 47(1): 109-123.

[74] Pirotte, A. 1999. "Convergence of the Static Estimation Toward the Long Run Effects of Dynamic Panel Data Models." *Economics Letters* 63:151-158.

[75] Rajagopal, D., S. E. Sexton, D. Roland-Holst, and D. Zilberman. 2007. "Challenge of Biofuel: Filling the Tank Without Emptying the Stomach?" *Environmental Research Letters* 2(4): 1-9.

[76] Rask, Kevin. 1995. "The Structure of Technology in Brazilian Sugarcane Production, 1975-87." *Journal of Applied Econometrics* 10(3): 221-232.

[77] Rees, Joseph V. 1994. *Hostages of Each Other: The Transformation of Nuclear Safety since Three Mile Island*. Chicago: University of Chicago Press.

[78] Roberts, Michael J. and Wolfram Schlenker. 2010. "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate." NBER Working Paper No. 15921.

[79] Robertson, G. Philip, Virginia H. Dale, Otto C. Doering, Steven P. Hamburg, Jerry M. Melillo, Michele M. Wander, William J. Parton, Paul R. Adler, Jacob N. Barney, Richard M. Cruse, Clifford S. Duke, Philip M. Fearnside, Ronald F. Follett, Holly K. Gibbs, Jose Goldemberg, David J. Mladenoff, Dennis Ojima, Michael W. Palmer, Andrew Sharpley, Linda Wallace, Kathleen C. Weathers, John A. Wiens, and Wallace W. Wilhelm. 2008. "Sustainable Biofuels Redux." *Science* 322(5898): 49-50.

[80] Rose, Nancy L. 1990. "Profitability and Product Quality: Economic Determinants of Airline Safety Performance." *Journal of Political Economy* 98(5): 944-964.

[81] Rosegrant, Mark W. 2008. "Biofuels and Grain Prices: Impacts and Policy Response." Testimony for the U.S. Senate Committee on Homeland Security and Governmental Affairs, May 7, 2008.

[82] Rothwell, Geoffrey S. 1989. "Stock Market Reaction to Nuclear Reactor Failures." *Contemporary Policy Issues* 7(3): 96-106.

[83] Runge, C. Ford and Benjamin Senauer. 2007. "How Biofuels Could Starve the Poor." *Foreign Affairs* 86(3): 41-53.

[84] Rust, John and Geoffrey Rothwell. 1995. "Optimal Response to a Shift in Regulatory Regime: the Case of the US Nuclear Power Industry." *Journal of Applied Econometrics* 10(S1): S75-S118.

[85] Searchinger, Timothy, Ralph Heimlich, R.A. Houghton, Fengxia Dong, Amani Elobeid, Jacinto Fabiosa, Simla Tokgoz, Dermot Hayes, and Tun-Hsiang Yu. 2008. "Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change." *Science* 319(5867): 1238-1240.

[86] Sexton, Steven, Deepak Rajagopal, David Zilberman, and Gal Hochman. 2008. "Food Versus Fuel: How Biofuels Make Food More Costly and Gasoline Cheaper." *Agricultural and Resource Economics Update* 12(1): 1-6.

[87] Stock, James H. and Mark W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97(460): 1167-1179.

[88] Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20(4): 518-529.

[89] United States Department of Agriculture (UDSA). 2007. *Crop Calendars.* http://www.usda.gov/oce/weather/CropCalendars/index.htm.

[90] Verma, Kiran, Barry M. Mitnick, and Alfred A. Marcus. 1999. "Making Incentive Systems Work: Incentive Regulation in the Nuclear Power Industry." *Journal of Public Administration Research and Theory* 9(3): 395-436.

[91] White, Matthew W. 1996. "Power Struggles: Explaining Deregulatory Reforms in Electricity Markets." *Brookings Papers: Microeconomics* 201-250.

[92] Zhang, Fan. 2007. "Does Electricity Restructuring Work? Evidence from the U.S. Nuclear Energy Industry" *Journal of Industrial Economics* 55(3): 397-418.

[93] Zhang, Zibin, Dmitry Vedenov, and Michael Wetzstein. 2007. "Can the U.S. Ethanol Industry Compete in the Alternative Fuels Market?" *Agricultural Economics* 37(1): 105-112.