Three Essays in Energy and Environmental Economics

by

Paul A. Brehm

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2017

Doctoral Committee:

Associate Professor Hoyt Bleakley, Co-Chair
Professor Ryan Kellogg, University of Chicago, Co-Chair
Assistant Professor Catherine Hausman
Assistant Professor Shaun McRae, ITAM

Paul A. Brehm
pabrehm@umich.edu
ORCID iD: 0000-0001-5244-9297

ACKNOWLEDGMENTS

I thank my wife, Maggie, my parents, Ellen and Alex, my brother, David, friends, co-authors, and dissertation committee. Without their help, support, and guidance this dissertation would not have been possible. I am grateful for research funding provided by Rackham Graduate School and MITRE.

For helpful comments and discussions, I thank Ryan Kellogg, Hoyt Bleakley, Catie Hausman, Shaun McRae, Eric Lewis, Anirudh Jayanti, Andrew Usher, Dan Ackerberg, Ying Fan, Maggie Brehm, Alecia Cassidy, Traviss Cassidy, Alan Griffith, Kookyoung Han, Julian Hsu, Aristos Hudson, Sarah Johnston, Andrew Litten, Chris Sullivan, Yiyuan Zhang, Soren Anderson, Jeremy West, Akshaya Jha, Arthur van Benthem, Kelsey Jack, Erin Mansur, Berkeley's Energy Camp attendees, Heartland Workshop attendees, CU Environmental and Resource Economics Workshop attendees, and seminar participants at Michigan State University and the University of Michigan. For some Texas data, I thank Reid Dorsey-Palmateer. All errors are my own.

Contents

List of Tables

List of Figures

List of Appendices

ABSTRACT

Chapter 1: We exploit a federal oil lease lottery to examine how markets correct for initial misallocation. Lottery participants included oil companies, as well as individuals without the capital or expertise to drill for oil. In the absence of reallocation, we expect less drilling on leases won by individuals. We find that leases won by firms and individuals have similar short- and long-term outcomes, suggesting that secondary markets rapidly and efficiently correct for misallocation to individuals. However, the small subset of parcels with nearby oil production have 50% *less* drilling when they are won by firms. We develop a simple model to demonstrate how information asymmetry adversely affects firms to a greater degree. Because individuals have larger gains from trade, they are less likely to have their decision to trade affected by asymmetric information and are more likely to trade with a nearby producing firm.

Chapter 2: Between 2007 and 2013 the natural gas price dramatically declined, in large part due to hydraulic fracturing. Lower natural gas prices induced switching from coal generation to natural gas generation; I find 2013 carbon emissions fell by 14,700 tons/hour as a result. Lower prices also incentivized new investment in natural gas capacity. The more efficient capital stock led to an additional decrease of 2,100 tons/hour in 2013. I estimate 65-85% of this new capacity was constructed because of lower gas prices. Using a social cost of carbon of $35/ton, I value the total decrease at roughly $5.1 billion.

Chapter 3: We examine the relationship between airfares and oil prices in the Australian airline industry. We find pass-through rates in excess of 100% that vary depending on the amount of competition on a route. We also find evidence that different types of products can have heterogeneous pass-through structures – pass-through rates on non-stop routes are more responsive to competition than on one-stop routes. Our results have important implications for environmental policy in industries with imperfect competition and differentiated products.

Chapter 1

To Trade or Not to Trade: Oil Leases, Information Asymmetry, and Coase (With Eric Lewis)

1.1 Introduction

The Coase Theorem (Coase, 1960) suggests that, in the absence of frictions, initial asset misallocation should be corrected through secondary market trading. Information asymmetry is one type of friction that can disrupt reallocation (Akerlof, 1970; Myerson and Satterthwaite, 1983; Hendricks and Porter, 1988). Our paper seeks to answer two related questions. First, in the absence of large frictions, can we find evidence that initially misallocated assets can quickly and efficiently be reallocated, as the Coase Theorem predicts? Second, does the presence of information asymmetry change our results?

We examine a 1970's oil (and gas) lease lottery that only cost $10 per lease to enter[1], allowing many different individuals and firms to enter the lottery. While winning oil firms had the capacity to exploit their leases, winning individuals likely lacked the necessary capital and expertise. Many of these individuals entered the lottery with the intent of flipping their winnings to firms and securing a quick profit, but were likely poor initial matches for the drilling rights. We hypothesize that in the absence of secondary markets, parcels won by firms would have seen much more development.

We test whether initial misallocation to individuals (instead of firms) as a result of the lottery led to short- and medium-run differences in drilling and production. While assignment via lottery provides exogenous variation conditional on entry, we also need to correct for potential bias from

[1] $10 in 1975-1978 is worth about $41 in 2015.

endogenous entry. Indeed, our data show that firms and individuals submit entries for different types of leases at different rates. We use two strategies to control for bias from endogenous entry.

Our first strategy uses the fact that each lease had three "winners", a first-place winner that actually claimed the lease, as well as two back-up winners. We examine a restricted set of parcels where exactly one of the three winners is a firm. Within this subsample, knowing the identity of the initial winner provides no information about the underlying characteristics of a lease. As a result, we can estimate an unbiased treatment effect from having a firm win a lease. Our second strategy uses the full set of parcels and relies on observables to control for bias from endogenous entry. Results are broadly consistent across our two specifications.

Consistent with the Coase Theorem, we find that leases without information asymmetry are likely to be reallocated, resulting in similar ex-post probabilities of drilling and production for leases won by firms and those won by individuals. This finding provides evidence in favor of the efficiency of secondary markets that contrasts with a recent literature (Bleakley and Ferrie, 2014; Akee, 2009). Their papers show that correcting for initial market mismatches takes anywhere from 20 to 100 years, while our paper shows that reallocation can allow outcomes to be very similar in the immediate term.

Our findings likely differ from the existing literature because of the relative ease of correcting for initial misallocation in our setting. Underlying lease characteristics in our sample started out identical, in expectation; the previous literature examined convergence when assets started out with different underlying characteristics. Additionally, the deep pool of interested parties, reliable (if limited) information about leases, and the relative ease of contacting lessees all likely aided secondary market reallocation.

A subset of leases in our sample have nearby oil production. This nearby oil production provides private information to the nearby producer about the quality of their land, as well as the quality of nearby land. Private information creates a lemons problem and makes it difficult for parties to agree on a fair transfer price. Indeed, Wiggins and Libecap (1985) find that information asymmetry from nearby production impedes unitization on oil fields that are being developed because parties are unable to agree on how to share oil production.

We find that leases with nearby production have about twice as much drilling and production if they are initially allocated to individuals, as opposed to firms. This result is likely caused by individuals' greater propensity to reassign their leases to nearby producing firms that can make the best decisions about whether to drill on the land. Analysis of reassignment patterns confirms that firms are less likely than individuals to trade with nearby producing firms.

In order to more formally understand our results, we develop a simple theoretical model of lease reassignment and drilling that is based in part on Myerson (1985); Samuelson (1985). Information asymmetry in the model works to prevent trade because uninformed parties know that informed parties will only agree to a trade if the latter are paying a price at or below the fair market value – there is limited upside to engaging in trade for the uninformed party. The intuition of the model is that gains from trading with a nearby producing firm are larger for individuals, causing them to be more likely to overcome information asymmetry. Because nearby producing firms have information about the quality of the lease, they are better able to identify the best leases to drill. On the other hand, firms that retain leases they initially win are unable to internalize the nearby firms' information about their lease, causing them to be less likely to drill.

Our work improves the understanding of how (mineral lease) secondary markets reallocate assets. This is particularly relevant today because current fracking hotspots have a similar structure to the Wyoming oil lease lottery. For example, oil deposits in the Bakken shale were allocated haphazardly, from the perspective of farmers who were not expecting to drill for oil. In addition to similarities in the allocation process, the reallocation process remains comparable - direct negotiations between landowners and firms interested in buying drilling rights.

We also contribute to the understanding how asymmetric information can impede otherwise efficient market design. In the absence of information asymmetry and in the presence of a robust secondary market, our work suggests that initial allocations may not be important. However, the presence of information asymmetry suggests that it is better to initially allocate an asset to somebody who cannot use it than it is to allocate the same asset to a random firm (if there is a robust secondary market). While the former allocation type is initially worse, secondary markets are better able to correct for this mismatch because the gains from trade are larger. This ties to theoretical work by Myerson and Satterthwaite (1983) showing that when the range of typical seller valuations does not overlap with the range of typical buyer valuations, efficient trade can occur that is both individually rational and incentive compatible—even though buyers and sellers still have uncertainty about each other's valuations. Our results demonstrate a novel way of avoiding information asymmetry problems – allocate assets to owners who must trade them.[2]

Finally, we demonstrate that under certain conditions (information symmetry and a robust secondary market), lotteries can be a relatively efficient way of allocating public resources. In addition to oil and gas, many other natural resources (timber, coal, water, etc.), as well as spectrum, Inter-

[2]Our work also ties to a strand of corporate finance literature showing that information asymmetry problems can be mitigated by bank-imposed opacity (for example, Dang et al. (2017) show that bank-imposed opacity promotes asset liquidity).

net domain names, and carbon permits are public resources that could be allocated via lottery.[3] Many economists would suggest that an auction is a superior method of allocating resources. Our results suggest that the efficiency cost of lotteries is lower than expected in some settings. Indeed, lotteries with a robust secondary market could be a preferable form of allocation if they are less costly to implement than an auction, if they are seen as a more equitable way of distributing public resources, or if they generate higher revenues.[4]

Our paper proceeds as follows. Section 2.2 discusses background for our empirical setting, including the lottery system and leasing rules. Section 1.3 describes our data. Section 1.4 explains our empirical specification and main results. Section 1.5 develops our model and Section 2.9 discusses and concludes.

1.2 Background: Oil Lottery and Reallocation

This paper focuses on oil and gas leasing, drilling, and production activity on United States Federal Government owned land. The Bureau of Land Management (BLM), a division of the Department of the Interior, allocates leases to drill for and produce oil and gas. We focus our study on Wyoming, where approximately 52% of the land is owned by the federal government, and where there has been significant oil and gas production (Fairfax and Yale, 1987).

Our analysis uses leases that were allocated by the BLM via lottery, providing randomization that is key to our identification strategy. The BLM began using a lottery system for "non-competitive" land in 1960 as a way to provide an orderly and fair allocation (Bureau of Land Management, 1983).[5] "Non-competitive" land was designated as such if it was on the site of a lapsed lease, the previous lease was not known to be productive, and the land was at least one mile away from known oil or gas production (Fairfax and Yale, 1987).[6] The lottery system ended in

[3] One example with a similar reallocation process is highlighted in a recent Wall Street Journal article highlighting how 30 of 104 entrants into a wireless spectrum auction were individuals taking advantage of a 25% discount designed to encourage new entrants. Successful bidders who qualify for discounts are allowed to lease all their airwaves to larger companies and sell the airwaves after 5 years.

[4] Many state-run lotteries (like "Powerball") generate more revenue than the value of the prize. Additionally, auctions can generate substantially less revenue than the expected value of the prize if competition is limited.

[5] "Before 1960, these tracts were offered on a first-come, first-served basis. When particularly promising tracts were due to be posted as available, long lines formed at the land offices. Fights often broke out, disrupting business and, in some instances, injuring employees trying to control the crowds. The simultaneous oil and gas lease drawing was developed to establish an orderly and fair system of awarding these noncompetitive leases." (Bureau of Land Management, 1983).

[6] For land closer to known production, the government used an auction. For land that had never been leased, it used a first-come first-serve system.

1987 when the BLM switched to using auctions for this type of land.[7,8]

Lottery specifics remained constant during our sample period (1975 through 1978), though they changed slightly during later lotteries. Each month, regional BLM offices would compile a list of the parcels that would be offered in the lottery.[9] Interested individuals and firms typically had about a week to submit an entry card to the regional BLM office for each lottery they wanted to enter (Bureau of Land Management, 1983). Many entrants filed multiple cards, one per plot for multiple plots, with a $10 filing fee for each entry card submitted. As this entry fee was very low, it was unlikely to restrict individuals who had less industry experience and capital from entering.

The regional BLM office would then draw three entry cards – one for the first-place winner, along with two runner-ups – which we refer to as the first-, second-, and third-place winners. The holder of the first place winning entry card had 30 days to submit a rental payment, equal to $1 per acre, in order to secure the lease. If the first-place winner did not respond with a rental payment within 30 days, the BLM contacted the second-place winner with a notification that the first-place winner had not responded and that the second-place winner was now eligible to submit the rental payment and obtain the lease. In the case of no response, the BLM then turned to the third-place winner.[10] Lottery losers were informed of their loss by the return of their entrance cards (Bureau of Land Management, 1983).

In order to retain leases after winning, new lessees were required to continue to comply with BLM leasing rules. Lessees paid a rental fee of $1 per acre per year in the years prior to any oil or gas production. After leases began production they paid a royalty, typically 12.5 percent, on revenues from production. If a lease had oil or gas production within 10 years of the lease start date, the lease continued until production ended. If there was no production within 10 years and if qualifying drilling operations were not in progress, leases generally expired and were returned to the BLM. A lease also ended if the lessee formally abandoned it or failed to pay rental fees (Fairfax and Yale, 1987). These expired and abandoned leases were then re-offered through lottery–or via auction, if they expired after 1987.

Two institutional features helped leases to be reallocated to the optimal owners. First, there

[7] One of the main reasons for this change was to combat middlemen from "filing services" who charged excessive fees to file entry cards on behalf of unsophisticated parties. For example, a $250 filing fee might be charged to a retiree in Florida (with the middleman keeping $240 and sending the $10 fee in to the BLM).

[8] The BLM still uses lotteries for parcels that receive no bids in the auction.

[9] The Cheyenne office's region included Wyoming, as well as small parts of Nebraska, Kansas, etc. We exclude leases outside of Wyoming due to data limitations.

[10] By matching winner records with transactions records, we estimate that approximately 98% of leases were claimed by the first-place winner.

were low search costs: the BLM maintains an open records office where anyone can easily look up the name and contact information of current lessees. Low search costs increase the number of potential buyers who might contact a lessee. More potential buyers may increase buyer willingness to sell both by raising the price via competition, as well as by giving the individual additional signals about the expected value of the lease.

Additionally, misallocation is easily observed. When firms visit the BLM open records office, it is easy for them to determine whether a current lessee is an oil and gas firm or an individual. Firms were very aware that many individuals entered the lotteries hoping to strike it rich, which they could likely only do if they transferred their lease to a firm. In these cases, the individuals' valuations of leases in the absence of transfer was likely very low, and possibly negative, as the lease required paying annual rental fees. In contrast, firms typically had positive valuations.

1.3 Data Description & Summary Statistics

1.3.1 Data Description

Our lottery sample is compiled from Bureau of Land Management (BLM) lease data. Lotteries were held monthly from January 1975 to December 1978, with roughly 225 parcels for lease per month. We drop a small number of parcels ($<1\%$) due to illegibility in the raw data or for having fewer than three winners. This leaves a total sample of 10,762 parcels offered over 48 months. The data include information on the boundaries of the leases, allowing us to match the leases to drilling and production activity. The data also include the total number of entry cards submitted for each parcel to be leased, as well as the names of the first-, second-, and third-place winners. Appendix A.1 discusses this data in greater detail.

The winners' names allow us to determine whether the winner was a firm or an individual. We first search the names for terms like "Inc.", "Production", or "Corporation." We then reviewed each lease's winners by hand to ensure that they were accurately categorized. Most firms appear to be oil and gas production companies. Some individuals only appear once or twice and appear to not have connections to the oil industry. Other individuals appear to be more "sophisticated", winning frequently and having close connections with the oil industry.[11]

Our primary outcomes of interest, drilling and production activity, use data from the Wyoming Oil and Gas Conservation Commission. These data have been lightly edited to improve quality by the United States Geological Survey (Biewick, 2011). It includes both the date and Public

[11] We discuss individuals entering on behalf of firms and other potential manipulations in Appendix A.1.

Land Survey System location for each well drilled within Wyoming. This allows us to determine whether a lease had nearby production when it was listed in the lottery, as well as whether drilling or production occurred on the lease itself.[12]

We define nearby production to be an oil-producing well drilled within five years of a lease's start and located within 2.6 miles of the lease.[13] In Appendix A.4 we consider a variety of alternative definitions of nearby production and find that our results are broadly consistent.

We also compile lease transfer outcome data from the BLM's LR2000, an administrative database of federal leases. The LR2000 includes detailed information, including leases transfers, lease terminations, and lease partners. This dataset allows us to measure how quickly markets correct for misallocation. One limitation of the LR2000 is that a selection of leases is not included because they expired prior to the digitization of their data. However, we find that approximately 95% of leases from our lottery data can be linked to the LR2000. Unfortunately, increasing this match rate to 100% is not possible as paper records of leases and lease transfers are destroyed fifteen years after lease expiration.[14]

Table 1.1 shows summary statistics for the parcels that were won through the lottery. Columns (1) and (2) summarize our full sample.[15] A typical parcel had about 600 entries submitted and an area just over one square mile (640 acres), although the variance was large. Wyoming well spacing rules imply a maximum of about sixteen oil wells (or four natural gas wells) on the average parcel.[16]

[12]Production data are only available starting in 1978. The fact that our first leases begin in 1975 but our first production does not appear until 1978 could be problematic. If, for example, a well was drilled in 1975 and production finished before 1978 we would be unable to capture this production. However, this is very unlikely for two reasons. First, immediate drilling was rare because permits needed to be secured and infrastructure (such as roads) frequently needed to be built before drilling could occur. Second, even if drilling happened within 3 years of the start date, most wells have long production lifespans, such that we would still observe the well producing by 1978. For example, we find that of all wells drilled in Wyoming between 1979 and 1985, less than 10% of wells have a lifespan less than 3 years.

[13]We exclude gas-producing wells because most participants in the colloquially known "oil lease lottery" were looking for oil, as it was much more profitable. In 1984, oil accounted for about 95% of the energy produced from hydrocarbons in Wyoming (Fairfax and Yale, 1987). Additionally, the Supreme Court (Phillips Petroleum Co. v. Wisconsin (347 U.S. 672)) ruled that wellhead natural gas prices were subject to regulation by the Federal Power Commission. Due to price ceilings on wellhead prices and the need to build expensive pipeline infrastructure to transport natural gas, wells producing natural gas were much less profitable and less likely to be an important source of information asymmetry.

[14]E-mail exchange with Wyoming BLM public room, Feb 13, 2015.

[15]Columns (3) and (4) summarize a restricted sample that has better identification. The restricted sample, where exactly one of the three winners was a firm, will be discussed in greater detail in Section 1.4.1).

[16]To reduce common pool problems, Wyoming restricts where wells can be located. Within each square mile section, land is divided into four square quarter-mile sections and further subdivided into sixteen sixteenth-of-a-square-mile squares known as quarter-quarters. Wyoming allows there to be one natural gas well per quarter, and the wells must be located at the center of the quarter. For oil wells, there can be only one well per quarter-quarter, and each well

We find that firms made up a relatively small fraction of winners. Only 6% of parcels had a firm as the first place winner. Similarly, 6% of all winners (first-, second-, or third-place) were firms. These leases typically had low expected productivity, only 3% resulted in oil production within twelve years; only 5% had production within thirty years.[17]

Table 1.1: Summary Statistics at the Parcel Level

	Full Sample		One Firm Winner	
	mean	st.dev.	mean	st.dev.
Number of entries	598.09	824.27	427.32	644.99
Area (sq. miles)	1.11	1.09	0.97	1.05
Number of firms among winners	0.19	0.42	1.00	0.00
Firm is first place winner	0.06	0.25	0.34	0.48
Nearby production indicator	0.22	0.42	0.21	0.41
Any drilling within 5 years	0.04	0.20	0.03	0.18
Any drilling within 12 years	0.08	0.28	0.07	0.26
Any drilling within 30 years	0.15	0.35	0.13	0.33
Any production within 5 years	0.02	0.12	0.01	0.11
Any production within 12 years	0.03	0.18	0.03	0.17
Any production within 30 years	0.05	0.23	0.04	0.21

Notes: The first two columns describe the entire sample (10,762 parcels). The second two columns describe the restricted sample where exactly one of the three winners was a corporation (1,800 parcels). The number of parcels is a round number (1,800) by chance, no sampling has occurred.

1.3.2 Ruling out Corruption & Testing for Randomness

Our identification strategy could be threatened if the lottery was not truly random – either because of corruption or other factors. If, for example, firms bribed BLM officials to win desirable parcels, that would invalidate our assumption that the underlying parcels were ex-ante identical. We now check whether there is evidence of lottery manipulation.[18]

We do not find that firms placed first a disproportionate amount of the time. Of the 2047

must similarly be located at the center of the quarter-quarter.

[17] Thirty years after the beginning of our leases, the original leases have generally lapsed and the lands are owned by winners of subsequent allocation procedures.

[18] We investigate this in part because of allegations of computer fraud surrounding the 1983 lottery of leases in the Amos Draw geological formation. A Department of the Interior study determined that earlier incidents of multiple winners were caused by entrants entering many parcels and that the frequency of multiple winners was within statistical expectations (referenced in a Chicago Tribune article). Nevertheless, we conduct this analysis to convince ourselves that the lottery was fair.

appearances by firms, 697 (34.0%) of these are in first place, 658 (32.1%) are in second, and 692 (33.8%) in third. While 34.0% of firms appearing in first place is slightly above the 33.3% that we would expect, the binomial distribution predicts that there is a 24% chance of observing 697 or more firms in first place, given 2047 total appearances. Within the restricted sample of 1,800 parcels we have a similar distribution of firms across winners (619 in first place), with a 16% chance that at least 619 firms appear in first place.

Table 1.2 tests whether parcels won by firms are similar on observables to those won by individuals. We restrict the sample to parcels where exactly one firm was amongst the three winners and compare parcel characteristics. Within this restricted sample, we find that parcels won by individuals and firms are very similar; those won by individuals have slightly more offers and are slightly larger, but the differences are well within the type of statistical variation we would expect to see. We therefore proceed under the assumption that the lottery was fair.

Table 1.2: Lease Comparison by Winner Type

	Individuals	Firms	Difference (p-value)
Offers Mean	430.99	420.33	0.74
Offers Variance	647.17	641.28	0.80
Acreage Mean	628.33	610.32	0.59
Acreage Variance	676.40	663.68	0.59
Nearby Production Mean	0.20	0.23	0.19
Nearby Production Variance	0.40	0.42	0.19

Notes: We restrict the sample to the 1,800 parcels where exactly one firm appeared amongst the three winners. Statistics for parcels won by individuals are reported in column (1), while column (2) reports those won by firms. Column (3) reports the p-value from an equality test.

1.4 Analysis

We first detail our empirical strategies and how we correct for endogenous entry. We then look at lease transactions data to see how quickly secondary markets transferred leases. Next, we examine whether there is evidence that the initial winner's identity affects drilling and production. Throughout our analysis we compare firms with individual winners.

1.4.1 Empirical Strategy

Primary Specification

A simple comparison of leases won by firms with leases won by individuals will not correct for lottery participants endogenously choosing which lotteries to enter. Table 1.3 shows that the total number of entries for a given lease is negatively correlated with the probability that the winner is a firm. It also shows that the number of entries is positively correlated with both ex-ante (total acreage) and ex-post (probability of drilling) measures of profitability. That is, individuals tend to crowd out firms on parcels with higher expected productivity. As a result, firms are less likely to win leases that eventually had drilling or production.

Table 1.3: Variable Correlations: Full Sample

	1	2	3	4	5	6	7
1: Number of entries	1.00						
2: Area	0.41	1.00					
3: Nearby production	0.14	-0.14	1.00				
4: Drilling within 12 years	0.20	0.08	0.18	1.00			
5: Production within 12 years	0.17	0.02	0.16	0.63	1.00		
6: Number of winning firms	-0.10	-0.06	-0.02	-0.03	-0.02	1.00	
7: Firm is first place winner	-0.06	-0.04	0.00	-0.02	-0.02	0.58	1.00

Notes: Correlations between selected variables for each observation using the unrestricted sample with all 10,762 parcels.

To illustrate the effects of a naïve comparison, suppose we ran the following regression where the probability of drilling at time t (Y_i) is regressed on a constant and an indicator for whether the winner was a firm (F_i):

$$Y_i = \alpha_0 + \beta_1 F_i + \epsilon_i \tag{1.1}$$

In this specification, β_1 estimates the naïve treatment effect of a firm winning a lease: it will be biased because ϵ_i is correlated with F_i. Our specification would not control for the fact that plots won by firms are disproportionately low quality.

Our primary specification breaks the correlation between ϵ_i and F_i by restricting our sample to the set of leases where exactly one firm was among the first-, second-, and third-place winners. Within this subsample, a firm is no more likely to be in first place than they are to be in second- or third-place. Therefore, the treatment effect of a firm winning a lease will be uncorrelated with unobservables and we obtain an unbiased estimate. All differences in outcomes are therefore attributable to differential behavior by different classes of winners.[19]

The parcels in this subsample are similar to the total sample, though of slightly lower quality. Parcels in the subsample tend to have fewer entrants on average (425 versus 598 in the full sample), slightly smaller acreage (0.97 square miles versus 1.11 in the full sample), and lower likelihoods of eventual drilling and production. This is consistent with the fact that individuals crowd out firms on leases with higher expected profitability. Table 1.4 shows correlations between variables within the subsample – it is similar to Table 1.3, which looks at the entire sample. However, the type of winner that is in first place is no longer correlated with the number of offers or the parcel size.

Table 1.4: Variable Correlations: Restricted Sample

	1	2	3	4	5	6	7
1: Number of entries	1.00						
2: Area	0.40	1.00					
3: Nearby production	0.09	-0.15	1.00				
4: Drilling within 12 years	0.20	0.09	0.16	1.00			
5: Production within 12 years	0.17	0.03	0.16	0.63	1.00		
6: Number of winning firms	.	.	.	.	.	.	
7: Firm is first place winner	-0.01	-0.01	0.03	-0.00	-0.02	.	1.00

Notes: Correlations between selected variables for each observation using the restricted sample with exactly one firm winner (1,800 parcels).

Our primary specification using this subsample is:

$$Y_i = \alpha_0 + \beta_1 F_i + \beta_2 NearbyProd_i + \beta_3 NearbyProd_i * F_i + \Omega X_i + \epsilon_i \tag{1.2}$$

Our primary specification includes an indicator for nearby production, as well as an interaction term between the identity of the lessee and the presence of nearby production. We also include control variables (a spline in the number of offers, total acreage, and month-of-lottery fixed effects) that are not necessary but may improve precision.[20,21]

In this specification, β_1 informs us about the effect of assigning a lease to an individual or a firm. This coefficient will help us to answer our first research question – in the absence of large

[19]We would like to examine leases where exactly two of the three winners are firms. However, the very small sample size of this exercise prevents us from making meaningful inference.

[20]We use a linear probability model, which is appropriate for calculating average treatment effects (Angrist and Pischke, 2008). The AIC and BIC suggest a linear probability model is a better fit for our data than a logit or probit.

[21]Results without the control variables are similar.

frictions, can we find evidence that initial asset allocations do not matter? A large and positive β_1 will tell us that leases won by individuals are not being reassigned to optimal owners - and that leases won by firms have more drilling and production.

β_2 informs us about the overall effect of having nearby production on a lease. Note that this will capture several effects. Because nearby oil and gas increases the chances of oil and gas on one's own land, we expect that drilling and production will be more common when there is nearby production. However, this effect may be diminished to the extent that information asymmetry from nearby production inhibits optimal asset reallocation for both firms and individuals.

β_3 captures the differential effect of nearby production on firms versus individuals. Because firms are in a better initial position to exploit leases that they win, we might expect this coefficient to be positive. To the extent that this interaction effect is negative, it must be because individuals are better able to reassign their parcels to firms with the highest value.

We analyze several outcome variables (Y_i) of interest; we look at lease reassignment by time *t*, drilling by time *t*, production by time *t*, probability of production given drilling by time *t*, days to well completion, and total production from producing wells.

Throughout our analysis, we use (Conley, 1999) spatial standard errors. We allow ϵ_i and ϵ_j to be correlated if section *i* and section *j* are within twenty miles of each other.[22]

Secondary Specification

We also consider a specification using our entire sample, which increases the precision of our estimates, but relies on the assumption that control variables eliminate bias from endogenous entry.

We rely on the number of offers for a lease, the total acreage of a lease, and a set of month-of-lottery fixed effects to control for unobservables that are correlated with the probability a firm wins a lease. The number of offers is highly correlated with ex-post measures of profitability like drilling and production. This is because entrants know that not all parcels are equal – they know the exact location of each parcel, which gives them important information about how far away the nearest road is, limited geological information, and the ability to find out limited information about the productivity of nearby plots. The least desirable parcels received less than ten entries, while the most desirable received several thousand. This kind of variance is not random and is possible because people have different expectations about the profitability of different plots.[23]

[22]Lewis (2015) shows that information about parcels in section *j*, more than 20 miles away from section *i*, do not provide information about how productive section *i* will be.

[23]Indeed, under certain conditions the ex-ante expected value of a parcel can be exactly determined by the number

Our secondary specification is run on the full sample of leases:

$$Y_i = \alpha_0 + \beta_1 F_i + \beta_2 NearbyProd_i + \beta_3 NearbyProd_i * F_i \\ + \beta_4 TotalAcreage_i + s(NumOffers) + \delta M_i + \epsilon_i \quad (1.3)$$

We use a cubic spline, s(), with six knot points to allow for flexibility in the relationship between outcome variables and the number of offers. M_i are month-of-lottery fixed effects. Our coefficients of interest remain β_1 and β_3; their interpretation is the same as in our primary specification. Results using our second specification are very similar to results using our primary specification.

1.4.2 Lease Transactions: Results

We first examine lease transactions in order to understand how firms and individuals treated their drilling rights. This will help to inform expectations about whether they are looking to drill on the parcels they have won, or perhaps are looking to flip the parcels for a quick profit.[24] Figure 1.1 uses our restricted sample with exactly one firm winner and displays how long it took for leases to be transferred for the first time.[25,26]

We find that lease transfers were an important part of lease development. Nearly 30% of leases won by individuals were transferred within the first year of a lease. In contrast, firms were 22% less likely to transfer leases during the first year. Instead, firms were the most likely to hold on

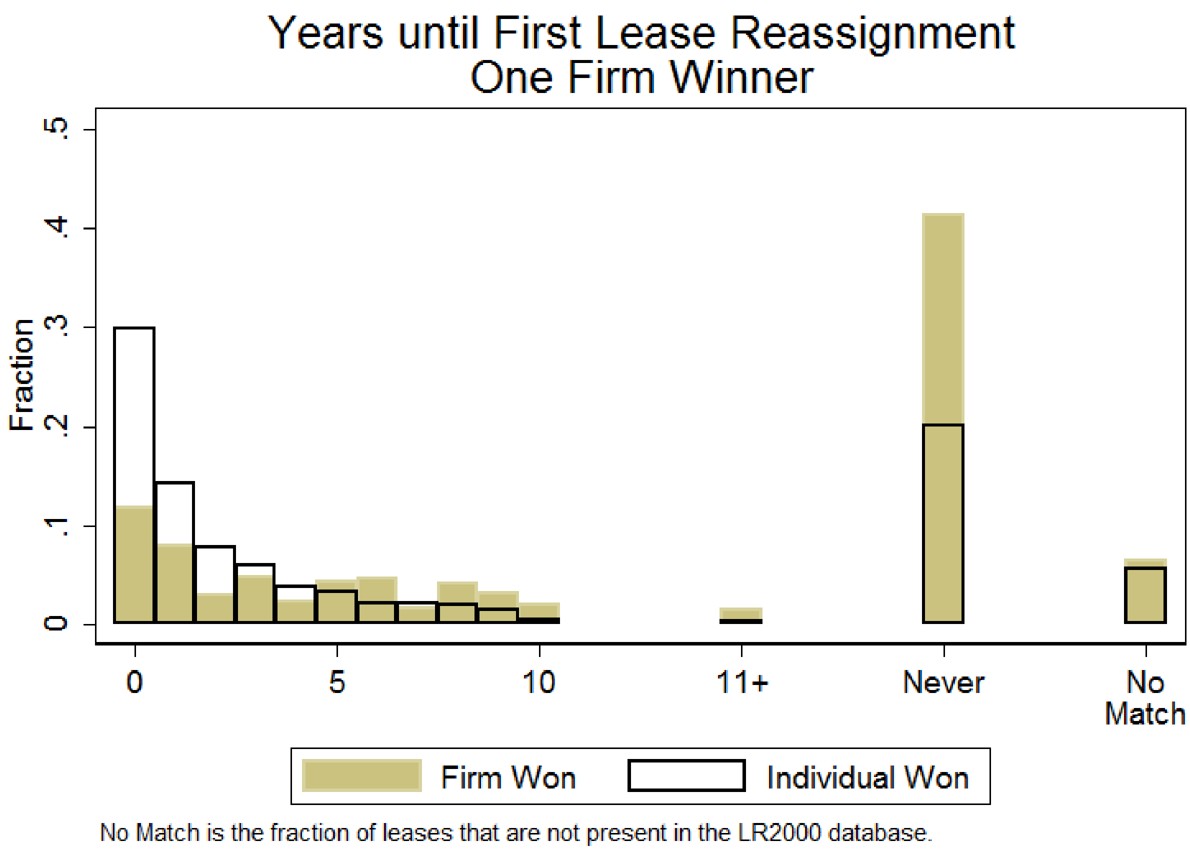
of entries. Suppose, for example, that the value of winning a plot is $1,000. The cost of entering the lottery is $10. Therefore, the expected value of entering a lottery is 1/N * $1,000 - $10. One would want to enter the lottery for a plot if they expected N to be less than 100 and would not want to enter if they expected more than 100 entries. If there were 200 potential entrants, one potential Nash equilibrium of an entry game involves all participants flipping a coin to determine if they enter. The expected number of entrants will be 100 and no party will find it profitable to deviate from the coin-flip entry strategy.

[24]One concern with the LR2000 is that the digitization of lease records began in the late 1980's; however, we are able to match over 95% of the leases because paper records were transcribed during that time. Importantly, we do not find a substantial difference in match rates between individuals and corporations. We believe that leases that do not appear in the LR2000 are most likely to have been passed over by the winner or to have expired early, though we cannot conclusively rule out other possibilities.

[25]Note that not all "first transfers" were to corporations. Indeed, our examination of the administrative database indicates many winning individuals first transferred portions of their leases to family members or other individuals. Many of these leases were re-transferred at later points in time, frequently to firms. We focus on "first transfers" because the administrative database only identifies parties by shorthand. Interpreting the type of party (firm vs. individual) is straightforward on a case-by-case basis. However, doing so systematically would be a major undertaking and almost certainly would not change our conclusions.

[26]Appendix A.2 presents regression results presents regressions using Equation 1.2 and our restricted sample to analyze years until initial lease reassignment.

Figure 1.1: Years until First Lease Reassignment



Notes: Histogram of the amount of time, in years, until a lease is transferred. Graph limited to leases where exactly one firm appeared among the first-, second-, and third-place winners. Individuals are in black outlines and firms are in green. Leases that do not have a recorded transfer are in the "Never" category, while leases that could not be linked to the LR2000 data are in the "No Match" category.

to the lease without *ever* transferring it —firms did this 22% more frequently than individuals. Therefore, the secondary market was an important mechanism to correct for misallocation and individuals were more likely to take advantage of the secondary market. Not only were a large fraction of leases initially allocated to individuals eventually transferred, but there was also a significant number of transfers from firms.

1.4.3 Firms vs. Individuals: Drilling Results

We first examine the total effect of having a firm in first place, regardless of whether a parcel has nearby production or not. We identify whether a lease has drilling within 1, 2, ..., and 15 years after the start of the lease, and whether it had production within 3, 4, ..., and 15 years of the start of the lease.[27] Leases without drilling activity generally ended after ten years, though some were extended because qualifying drilling operations were in progress. Due to the frequency of lease extensions, our preferred measure of 'final' differences due to initial allocation is at twelve years after the lease has begun. We also provide a limited expansion of this time frame to demonstrate that our results do not change substantially after lease expiration.

Figure 1.2 contains results of regressions analyzing the rollout of drilling activity using Equation 1.1. The dependent variable in panel (a) is an indicator for whether any drilling occurred on a lease up to X years following the lottery.[28] This regression is run for each of the first fifteen years following a lottery and the estimates are connected by a line. Panels (b) and (c) use an indicator variable for whether any production has occurred, with panel (c) restricting the sample only to parcels where drilling has occurred. This figure presents the total effect, including parcels with and without information asymmetry, of having a firm in first place using our restricted sample. We also include the p value of the test of equality.

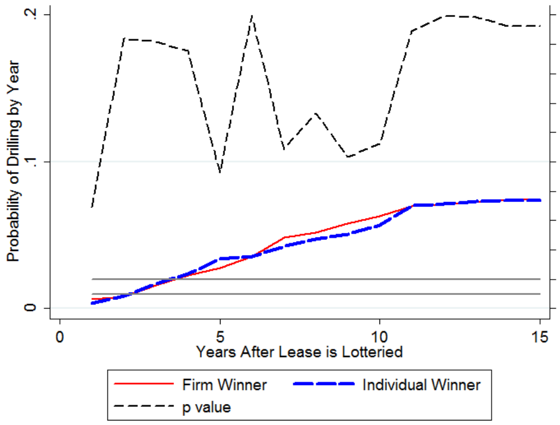
We find that few parcels are drilled immediately after the lease starts, but that by the twelfth year, about 7.5% of parcels experience drilling. We find that firms are slightly less likely to drill and produce, though the results are not statistically significant. Panel (c) suggests that firms may have slightly lower probabilities of production given drilling, though the differences are not generally significant and full sample results do not support this conclusion.

We also examine drilling and production outcomes over time using our full sample of 10,762 parcels. Figure 1.3 reports results without controlling for endogenous entry in the left column, while the right hand column includes variables like the number of offers for a lease to control for

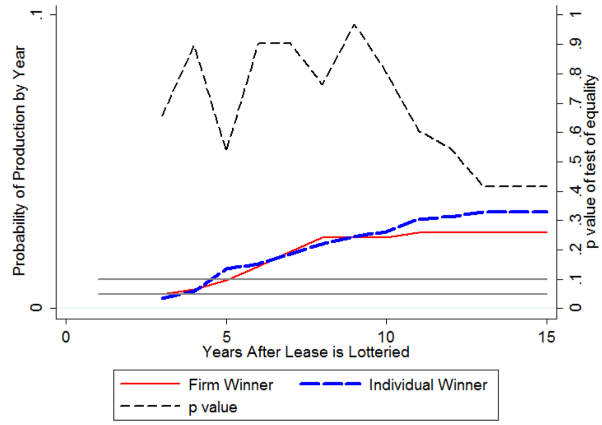
[27]We exclude production for the first few years because our lottery data begin in 1975 but our production data does not begin until 1978.

[28]For example, a lease in year six has had drilling if that drilling happened in any lease year from one through six.

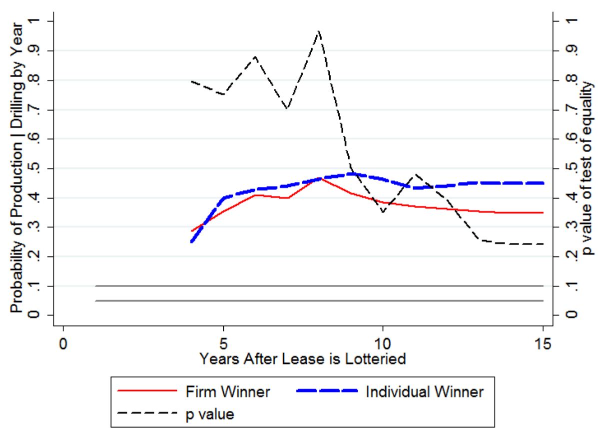
Figure 1.2: Restricted Sample Results: All Leases



(a) Drilling



(b) Production



(c) Production Given Drilling

Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm. This specification does not include an interaction between an indicator for nearby production and an indicator for whether the first place winner was a firm. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification.

land quality (detailed in Equation 1.3). The left column suggests that leases won by firms are less likely to have drilling or production. However, when we control for land quality on the right, most of the differences disappear. Indeed, results with controls look very similar to those in Figure 1.2. Point estimates suggest that firms may have lower likelihoods of drilling and production, though the results are not statistically significant. Overall, we find that initial assignment does not affect drilling and production outcomes.[29]

1.4.4 Nearby Production

We now investigate whether firms or individuals perform differentially better in the presence of nearby production. Table 1.5 presents the coefficients from regressions using Equation 1.2 results after the twelfth year of a lease (Table A.3 presents the full sample analog in Appendix A.3). The first row of Table 1.5 indicates that the presence of nearby production is associated with much higher rates of drilling and production (columns (3) and (4)). This is to be expected, as nearby oil is one of the best predictors of oil underfoot. Row 2 shows that leases won by firms and individuals, absent nearby production, have similar outcomes. It is similar to results in the previous section; we will briefly expand on this in the following section.

The third row, an interaction between nearby production and whether the initial winner was a firm, shows that leases won by firms have differentially lower rates of drilling and production than those won by individuals when nearby production is present. Figures 1.4 uses our restricted sample and presents results for each year following the lease lottery.[30] The dashed line shows the p-value from a test with the hypothesis that firms and individuals have the same results (for example, year twelve tests whether the sum of rows two and three from Table 1.5 is equivalent to zero).

Appendix A.4 presents a range of alternative specifications to check the robustness of our results. In particular, we consider specifications that vary the distance that we define as "nearby", the number of years that we define as "recent", and whether we include productive gas wells. Finally, we also present results including dry wells to demonstrate that nearby production is the important dimension, not nearby activity.

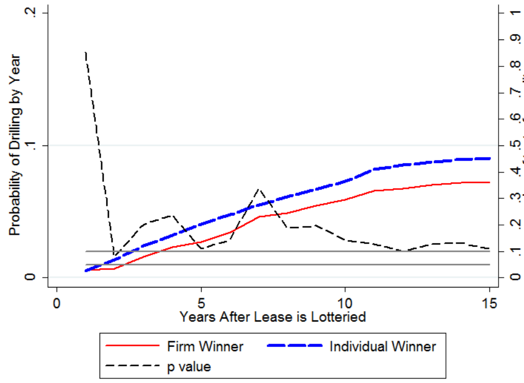
1.4.5 No Nearby Production

We now present results for parcels without nearby production; these results are similar to the total effect results presented in Section 1.4.3. Figure 1.5 uses our restricted sample and presents the

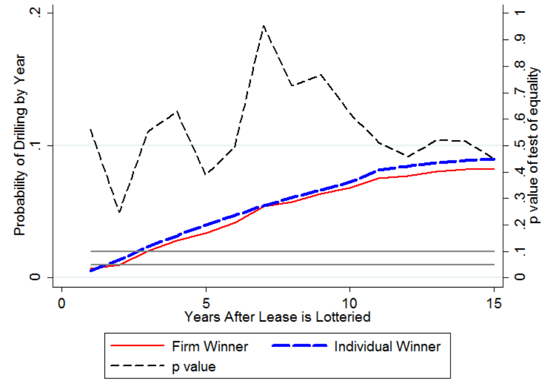
[29] Appendix A.5 looks at drilling time and production quantities. However, due to a lack of precision, we do not find these results to be substantive. They are included for completeness.

[30] Figure A.2 presents the full-sample analog in Appendix A.3

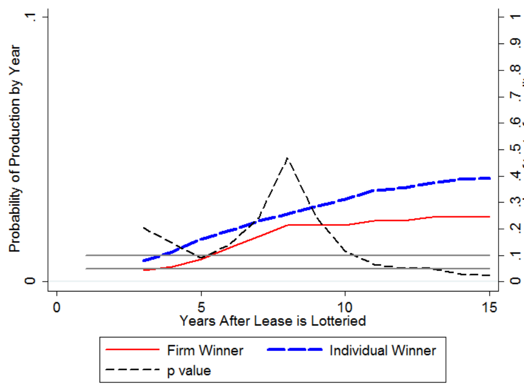
Figure 1.3: Full Sample Results: All Leases



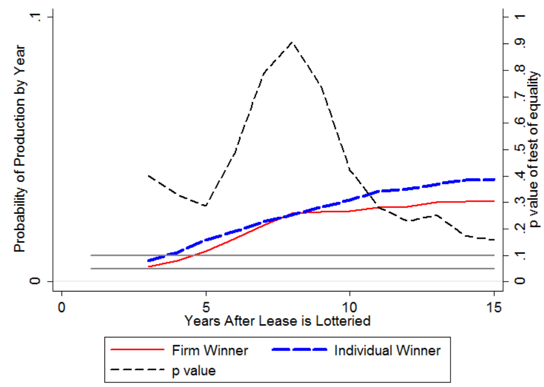
(a) Drilling, No Controls



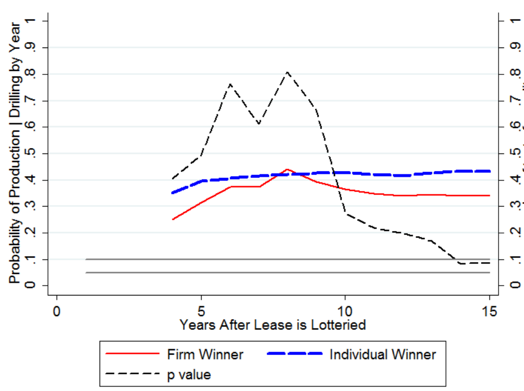
(b) Drilling, With Controls



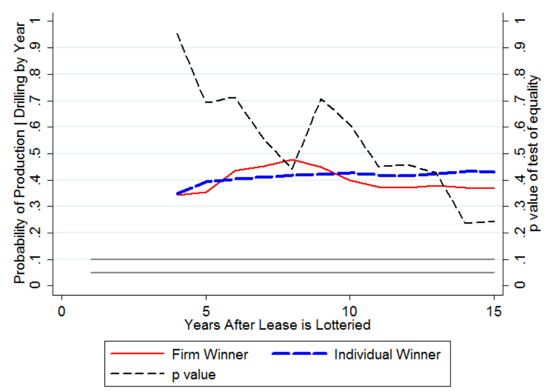
(c) Production, No Controls



(d) Production, With Controls



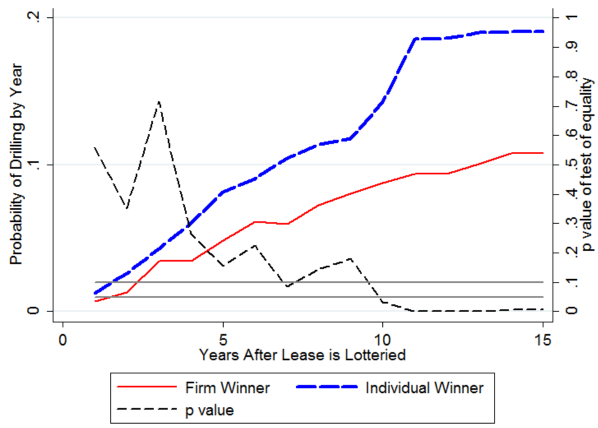
(e) Production Given Drilling,
No Controls



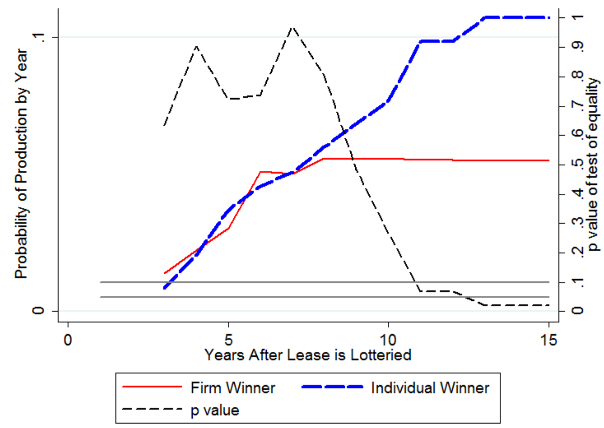
(f) Production Given Drilling,
With Controls

Notes: The left column presents the full sample with no controls. The right column controls for underlying land quality. The firm effect is the coefficient on an indicator for whether the first place winner was a firm. These specifications do not include an interaction between an indicator for nearby production and an indicator for whether the first place winner was a firm. The right vertical axis gives the p value of a test that the two means are not equal.

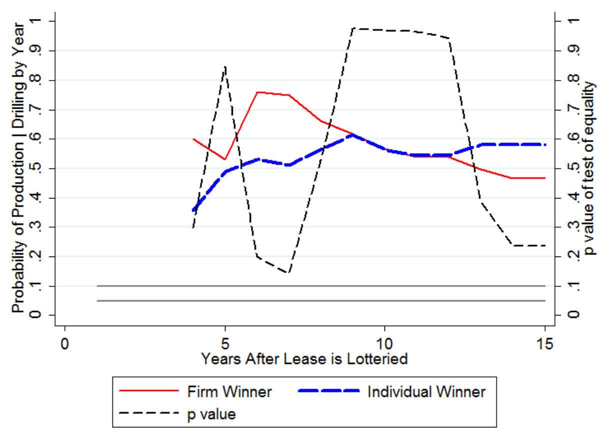
Figure 1.4: Restricted Sample Results: Leases with Nearby Production



(a) Drilling



(b) Production



(c) Production Given Drilling

Notes: Restricted sample with one firm winner. Estimates are reported for parcels with nearby production. The firm effect is the sum of the coefficient on an indicator for whether the first place winner was a firm, plus the coefficient on an interaction between the firm winner indicator and an indicator for nearby production. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification.

Table 1.5: Regression Results: Lease Transfers

	(1) Reassign Probability	(2) Log Time to Reassign	(3) Drill	(4) Prod	(5) P | D
Nearby Production Flag	0.027	0.019	0.136***	0.074***	0.130
	(0.038)	(0.146)	(0.026)	(0.020)	(0.149)
Firm Winner	-0.219***	0.880***	0.014	-0.000	-0.047
	(0.027)	(0.113)	(0.013)	(0.007)	(0.130)
Firm/Nearby Prod Interaction	-0.057	0.089	-0.106***	-0.043*	-0.269
	(0.067)	(0.233)	(0.031)	(0.024)	(0.222)
Offers & Acreage Controls	Yes	Yes	Yes	Yes	Yes
Lottery Fixed Effects	Yes	Yes	Yes	Yes	Yes
R squared	0.156	0.159	0.114	0.086	0.385
Observations	1800	1196	1800	1800	128

Notes: This table uses the probability of reassignment by twelve years (1), length of time given reassignment (2), probability of drilling by twelve years (3), probability of production by twelve years (4), and probability of production given drilling by twelve years (5) as dependent variables. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on. This table uses our restricted sample. Column (2) does not correct for selection into reassignment. Point estimates using a Heckman two-step are similar.

difference between firms and individuals on land without nearby production.[31,32] Results using both our primary specification on the restricted sample and our secondary specification on the full sample show that firms and individuals have similar outcomes when parcels do not have nearby production. Panel (c) shows that production given drilling is similar; the point estimates suggest that firms have slightly lower rates, though they are not significant. Table 1.5, Row 2, presents the coefficients for these results twelve years after the leases were lotteried.[33]

1.5 Reassignment and Drilling Model

The preceding sections present two key results. We find that information symmetry leads leases won by firms and individuals to have very similar outcomes, a result that is consistent with well-functioning secondary markets and provides evidence that the Coase Theorem has real-world applications. However, information asymmetry from nearby production imposes an important barrier to trade that causes leases won by firms to have *less* drilling and production. Because individuals are generally not drilling wells themselves, these drilling and production differences must arise due to different lease reassignment strategies.

We now develop a model to provide insight into our primary findings, with an emphasis on how leases won by firms can have *lower* rates of drilling than those won by individuals. Our model of asymmetric information is informed by Akerlof (1970); Myerson and Satterthwaite (1983) and Samuelson (1984). It is most analogous to results derived in Myerson (1985) and Samuelson (1985). One departure from the previous literature is that their models assume the seller is the party with private information. In our setting, the buyer will have superior information due to their nearby production.

1.5.1 Model Preliminaries

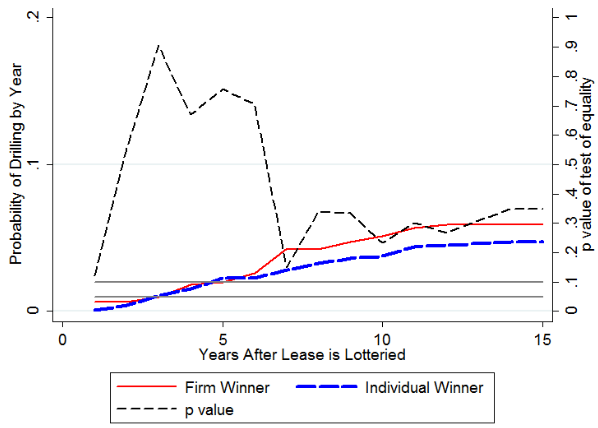
Our stylized model has three types of agents: firms that initially win a lease (F), individuals that initially win a lease (I), and the lowest cost firm that is interested in purchasing a lease (LCF). All parties are risk neutral. The lease has an underlying true value of production, $\theta \sim$ U[0,1]. In the base version of our model there is no information asymmetry and all agents can only form expectations about the true value of θ. In the absence of private information, E[θ] is the same for all parties.

[31] Figure A.3 presents the full-sample analog in Appendix A.3.

[32] Because less than 25% of our total sample has nearby production, our results are similar to those in Figure 1.2, which does not adjust for nearby production.

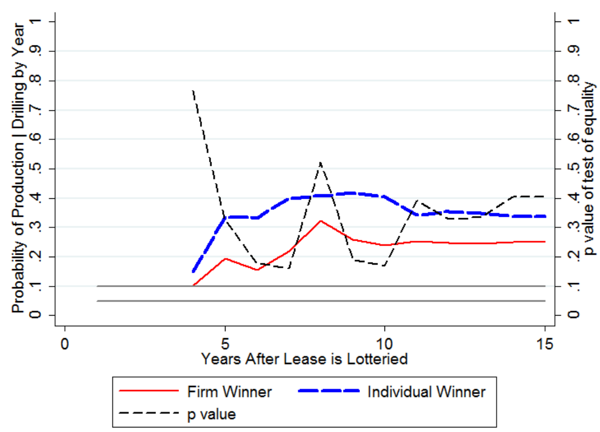
[33] Table A.3 presents the full sample analog in Appendix A.3.

Figure 1.5: Restricted Sample Results: Leases without Nearby Production



(a) Drilling

(b) Production



(c) Production Given Drilling

Notes: Restricted sample with one firm winner. Estimates are reported for parcels without nearby production. The firm effect is the coefficient on an indicator for whether the first place winner was a firm. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification.

Each type of lessee has a specific cost of drilling, $C_j, j \in \{I, F\}$, that is public knowledge and known with certainty. The LCF's cost of drilling, C_{LCF}, is also public knowledge and known with certainty. We assume that $C_I > C_F \geq C_{LCF}$. Note that the cost difference between firms and individuals is the only assumed difference between firms and individuals. It is reasonable for individuals to have the highest costs, as they generally do not have the capital and expertise to drill oil wells. Firms that are initial lessees and LCFs are likely to have similar costs when there is no nearby production.

The game proceeds in two stages. First, the lessee and LCF bargain over the lease. Bargaining happens via a take-it-or-leave-it offer, O_j, made by the LCF.[34] Following the reassignment stage, the lease owner decides whether or not drill. Expected payoffs from drilling equal the expected value of θ less the cost of drilling (C_j). Drilling happens if this value is positive.

1.5.2 Model Solution

The model solution without information asymmetry is quite simple. Because all information is public, parties do not learn anything as the bargaining process plays out. All parties will evaluate the expected value of drilling at 1/2. If $1/2 > C_{LCF}$ and $C_j > C_{LCF}$ then there are gains from trade. The LCF will look to buy the lease and drill on it. They will offer the minimum price such that the lessee will accept, the lease will be traded, and drilling will occur. If a lease is initially won by a firm, $1/2 > C_F$ and $C_F = C_{LCF}$, no trade will occur, but the initial lessee will make the same drilling decision that the LCF would.

Our setup yields predictions that are realized in the data. The eventual outcomes in terms of drilling and production will be similar, regardless of who wins the lease. Additionally, firms will be less likely to trade their lease than individuals. This is because the winning firm may be the lowest cost firm. We might also expect small transaction costs (not explicitly modeled) to inhibit trade between firms with similar drilling costs. We now turn to a richer model that explains our results on information asymmetry.

1.5.3 Information Asymmetry

Information asymmetry requires two adjustments to our assumptions. First, the LCF is replaced by a nearby production firm (NPF). The NPF is the potential buyer; their existing production

[34] In our basic framework, results are not sensitive to the type of bargaining framework. When we introduce information asymmetry, results will become sensitive to the choice of bargaining framework. This model demonstrates how information asymmetry *can* lead to worse outcomes for leases won by firms. We do not find that information asymmetry *always* leads to worse outcomes for leases won by firms.

allows them to know the true θ of the lease in question. Lease winners still do not have private information and can only form expectations about the value of θ. However, lessees are able to update these expectations based on offers they receive from the NPF. Notice that this will lead to a pooling equilibrium; a separating equilibrium is impossible because the NPF has no credible way to signal the true θ of the lease. Second, let costs be $C_I > C_F > C_{NPF}$. It is reasonable for the NPF to have the lowest costs, as they are already producing in the area.[35]

We now solve for the equilibrium of our simple bargaining model. The primary complication is that the lessee will update their expectations about the parcel's value as soon as an offer is made by the nearby producing firm.

If the initial lessee retains the lease after receiving an offer, O_j, it will receive their updated expectation of θ less their costs of drilling: $E[\theta|O_j] - C_j$. Thus, it will only trade if the offer is welfare improving: $O_j > E[\theta|O_j] - C_j$. Notice that there is no profitable deviation – if the initial lessee rejects an offer when $O_j > E[\theta|O_j] - C_j$ (or accepts an offer when $O_j < E[\theta|O_j] - C_j$), it will be worse off. $E[\theta|O_j]$ is determined by realizing that an offer of O_j indicates that the true value of θ lies somewhere between $O_j + C_{NPF}$ (adding C_{NPF} to account for the cost of drilling a well) and 1.[36] The uniform distribution means that $E[\theta|O_j] = \frac{(O_j + C_{NPF}+1)}{2}$.

Thus, the initial lessee will accept a trade if $O_j > \frac{(O_j + C_{NPF}+1)}{2} - C_j$, which reduces to:[37]

$$O_j > 1 + C_{NPF} - 2C_j \tag{1.4}$$

The NPF will be willing to offer a trade if the true value of θ is greater than its cost of acquisition plus the cost of drilling:

$$\theta - O_j - C_{NPF} > 0 \tag{1.5}$$

The NPF will find the minimum O_j such that the Equation 1.4 is satisfied and then offer this minimum O_j if it satisfies Equation 1.5. Substituting the first inequality into the second, we find that trade will happen when:[38]

$$\theta > 1 + 2C_{NPF} - 2C_j \tag{1.6}$$

[35] Due to the randomized nature of the lottery and the fact that most leases with nearby production had a mean (median) of more than 750 (325) offers, firms that win leases are unlikely to be the nearby producing firm.

[36] If an offer is made, this interval will always be of the form [X,1], where X is between 0 and 1. Neither O_j nor C_{NPF} can be less than 0; a lessee would never accept a negative offer and drilling costs are assumed to be positive. Additionally, the NPF will not make an offer such that $O_j + C_{NPF}$ is greater than 1, as this will result in a negative payoff for all possible values of θ.

[37] $O_j > \frac{(O_j + C_{NPF}+1)}{2} - C_j \Rightarrow 2O_j > O_j + C_{NPF} + 1 - 2C_j \Rightarrow O_j > 1 + C_{NPF} - 2C_j$.

[38] $\theta - (1 + C_{NPF} - 2C_j) - C_{NPF} > 0 \Rightarrow \theta - 1 - 2C_{NPF} + 2C_j > 0 \Rightarrow \theta > 1 + 2C_{NPF} - 2C_j$.

For all θs that satisfy Equation 1.6, the NPF will make the same bid, the minimum O_j that satisfies Equation 1.4.[39] Drilling will necessarily occur.[40] Notice that the individual's higher cost of drilling leads to a greater range of θs that are traded.

When trade does not occur, the lessee decides whether or not to drill based on whether $E[\theta|\text{no offer}] - C_j > 0$. In this situation, the lessee knows that θ lies between 0 and $1 + 2C_{NPF} - 2C_j$. Thus, $E[\theta|\text{no offer}] - C_j = \frac{0+1+2C_{NPF}-2C_j}{2} - C_j = \frac{1}{2} + C_{NPF} - 2C_j$. Notice that the expected gains from drilling are positive only when drilling costs are low ($C_j < \frac{1}{4} + \frac{C_{NPF}}{2}$). When drilling costs are high ($C_j > \frac{1}{4} + \frac{C_{NPF}}{2}$), the lessee will choose not to drill. Note that the firm will be in a better position to drill a retrained lease. However, individuals are more likely to trade leases, and traded leases are always drilled.

In equilibrium, drilling will happen whenever there is trade (when $\theta \in [1 + 2C_{NPF} - 2C_j, 1]$. If $\theta < 1+2C_{NPF} - 2C_j$, trade will not happen and drilling will only occur if $\frac{1}{2} + C_{NPF} - 2C_j > 0$.

1.5.4 Empirical Predictions

We are now able to examine our empirical findings. Is it possible for leases won by individuals to have *more* drilling than those won by firms? The answer depends on the costs of drilling.

Suppose that we are in an environment where the firm has low drilling costs and the individual lessee has high drilling costs ($C_I > \frac{1}{4} + \frac{C_{NPF}}{2} > C_F$). For example, C_{NPF} = 0, C_F = 0.1, and C_I = 0.5. If our data are actually represented by $\theta \sim$ U[0,1], our model predicts that trade and drilling will occur 90% of the time if individuals win all of the leases. In contrast, our model predicts trade will occur 20% of the time and drilling will occur 100% of the time when firms win. The higher level of drilling for parcels won by firms results from a high expected θ when there is no trade, combined with a low C_F.

On the other hand, consider when both firms and individuals have relatively high costs ($C_I > C_F > \frac{1}{4} + \frac{C_{NPF}}{2}$). For example, C_{NPF} = 0.5, C_F = 0.6, and C_I = 1.0. If individuals win all the leases, trade and drilling will occur 50% of the time. If firms win all of the leases, trade will occur 20% of the time and drilling will also occur 20% of the time. In this situation, assigning the leases to individuals will result in a higher proportion of them being drilled – consistent with our

[39] Again, note that there are no profitable deviations from this strategy. If the NPF offers more than the minimum O_j, it will not increase the range of θs that are accepted. If the NPF offers less than the minimum O_j, its offer will be rejected.

[40] We have not discussed what happens when θ is below the threshold. It is possible that the NPF could make an offer, knowing that it will not be accepted. This possibility can be rejected by either assigning a small transaction cost to making an offer, or by arguing that the NPF may not wish to provide information (in the form of an offer) to a competing nearby firm if it will not benefit from doing so.

empirical findings.

1.5.5 Testing Trade with Nearby Producing Firms

We now turn to examining the primary finding of our model: due to their lower drilling costs, firms will be less willing than individuals to trade their leases to nearby producing firms.[41] We scraped the identities of nearby producing well operators and matched these nearby operators to firms in the LR2000 lease history. We then estimate how frequently leases won by firms (relative to individuals) are traded with nearby producing firms using the following regression:

$$Trade_i^{NPF} = \alpha_0 + \beta_1 F_i + \Omega X_i + \epsilon_i \tag{1.7}$$

The dependent variable is whether we have matched a lessee as trading with a nearby producing well operator. The primary coefficient of interest, β_1, tells us whether firms are differentially less likely to trade with a nearby producing well operator. For this analysis, our sample is the set of leases with nearby production.

Our analysis is primarily limited because we cannot perfectly identify whether trade happened with the nearby producing firm. This leads to attenuation bias.[42] This imperfect match identification arises from four sources. First, our available well operator data does not also identify the nearby lessee. Any trades with a nearby lessee different from the operator will not be correctly matched.[43] Second, both the assignment data and the nearby operator data are incomplete. In particular, nearby operators are unavailable before 1978. Both datasets are likely missing observations, with greater completeness in later years. Third, neither dataset identifies corporate relationships. This means we will fail to identify a trade to a subsidiary or former employee of a nearby operator as a match. Finally, our analysis will be limited by imperfect identification of the "nearby" wells. For example, a well drilled outside of our five-year drilling window may have an owner that will be able to make the correct drilling decision, but we will not flag this type of

[41] Note that our earlier results demonstrate firms are less likely to transfer overall. This may not affect drilling frequency in the absence of information asymmetry because many firms are similar. However, failing to trade to nearby producers can have long-term consequences because they are likely to use their proprietary information to make different drilling decisions than the initial lessees.

[42] Because the dependent variable is binary, measurement error is non-classical and results in attenuation bias.

[43] In future work, we hope to incorporate some nearby owner data. Extracting the available nearby owners is a non-trivial task.

match.[44,45]

Table 1.6: Regression Results: Nearby Production Lease Transfers

	Sample with Nearby Firms		Sample with Distant Firms	
	(1)	(2)	(3)	(4)
Firm Winner	-0.079***	-0.077***	-0.023	-0.005
	(0.029)	(0.018)	(0.025)	(0.019)
intercept	0.049**	0.044**	0.027	0.000
	(0.024)	(0.021)	(0.024)	(0.015)
Full Sample	No	Yes	No	Yes
Offers & Acreage Controls	Yes	Yes	Yes	Yes
Lottery Year Fixed Effects	No	Yes	No	Yes
R squared	0.069	0.038	0.034	0.018
Observations	376	2399	443	2644

Notes: We examine how frequently leases are reassigned to firms that have nearby production. Column (1) uses our restricted sample and Column (2) uses the full sample. Columns (3) and (4) look at how frequently leases are reassigned to distant firms with production in the area. We allow reassignments to happen at any point during the lease. A pooled regression testing whether the 'Firm Winner' coefficient is different in columns (1) and (3) yields a p-value of 0.128. A pooled regression testing whether the 'Firm Winner' coefficient is different in columns (2) and (4) yields a p-value of 0.003.

Table 1.6 reports our primary results. In columns (1) and (2), we look at the set of leases with nearby production and examine whether leases are ever reassigned to a nearby producing firm. We find that firm winners about 7.5% less likely to have their parcels ever reassigned to a nearby

[44] Our analysis will also be limited because we are matching text strings. Abbreviations and typos introduce measurement error and lower our match rate. For example, "Gulf" can be matched with "Petrogulf Corporation" or "Gulf Oil". We use "Gulf" as the basis for matching because abbreviations are ubiquitous in the LR2000. Matching only with full names will cause many false negatives.

[45] Finally, we note that our preferred analysis looks at whether parcels are *ever* reassigned to the current nearby producing firm. In Appendix A.6 we examine whether parcels are traded to the current nearby operator before drilling has begun. We prefer our primary analysis because of the opaque nature of firm relationships. If nearby production is listed by "Conoco", it is possible that the nearby production was actually by Philips Oil Company, which was then bought by Conoco. However, we may not observe Philips as a nearby producer. Therefore, a lease traded to Philips, and subsequently to Conoco, would yield a false negative. The main drawback to this approach is that our analysis will occasionally identify matches not suggested by our model – where the trades happen after drilling. Fortunately, both approaches yield similar results as most matches are prior to drilling.

producing firms. Columns (3) and (4) use a sample without nearby production within 2.6 miles, but with production between 2.6 and 5.2 miles away from the lease. Here, we find that firms and individuals have similar rates of reassignment. These results show that when there is nearby production, firms are differentially less likely to reassign parcels to nearby producing firms.[46]

Is the magnitude of this effect large enough to explain the full drilling differences we find between leases won by firms and those won by individuals? By itself, no. However, it does demonstrate that firms are reassigning their leases differently than individuals. Additionally, we expect that attenuation bias from measurement error causes the coefficients to be smaller than their true values. These differing reassignment processes appear to be the primary cause of different lease outcomes.

1.6 Discussion and Conclusion

Overall, the accumulation of evidence suggests two facts. First, leases without information asymmetry are efficiently reallocated via a robust secondary market. The secondary market allows for similar drilling and production outcomes, regardless of initial allocation. Second, information asymmetry inhibits trade between firms and nearby producing firms. This leads to worse outcomes for these types of leases when they are initially assigned to firms. In our setting, we can conclude that it is better to initially have a bad match than it is to have a mediocre match. Because the mechanism – information asymmetry – is not specific to our setting, our findings are likely more broadly applicable.

Our finding that information asymmetry affected firms more severely than individuals may seem surprising at first. However, our theoretical model, combined with analysis of the transaction histories, reveals that our results are actually intuitive. Because the average firm is already a mediocre match for any given lease, the gains from trading that parcel are lower than the gains for an individual. As a result, many firms (in contrast to individuals) are insufficiently incentivized to trade their leases in the presence of information asymmetry.

One important policy implication is that lotteries can be a relatively efficient mechanism for allocating assets under certain conditions. For example, nature's "lottery" that allocated oil deposits

[46]Note that it is unclear whether we should expect the magnitude of the coefficients to be different between columns (1) and (3). On the one hand, firms are less likely to want to reassign to nearby producing firms than distant producing firms because it is likely that the nearby firms have a differential information advantage. On the other hand, distant producing firms may not have a cost advantage over the initial lessee. This would suggest that there are limited gains from trade. The fact that we see evidence of a differential suggests that there are reassignment gains due to cost differentials.

to farmers in the Bakken shale formation likely does not seriously impede the efficient outcome. Indeed, the Bakken's mineral rights reallocation process is very similar to the process in our setting; oil firms employ landmen to purchase the rights from individuals without the capital or expertise to develop the resources. Thus, our results provide evidence in favor of the United States' policy of allocating both surface rights and mineral rights to property owners.

Additionally, market designers should carefully evaluate whether information asymmetry is likely to be present. If it is, taking steps to ensure that assets are not allocated to intermediate-quality matches can yield tangible benefits. Potential applications include wireless spectrum auctions and electricity transmission markets. In electricity markets, for example, information asymmetry may arise due to proprietary information about generation costs. Our results suggest that assigning transmission rights to generating firms is inferior to auctioning the rights or assigning them randomly via lottery to individuals.

Our work suggests one primary avenue of further research. While lotteries can be an efficient way of allocating public resources, we do not yet understand how they affect government revenue. One of the government's primary goals when allocating a public resource is to raise funds.[47] We will compare lotteries and auctions run by the BLM to determine which allocation method raised more revenue. There are two possible ways for us to explore this question. We can match auctioned parcels with similar parcels that were lotteried at the same time. Additionally, we can compare revenues before and after the BLM switched to an auction for all parcels in 1987. Answering this question will give us a fuller picture of the relative welfare effects of auctions and lotteries.

[47] Li (2014) finds that auctions generate greater social welfare than lotteries for Chinese license plate distributions. However, China does not allow for reallocation of licenses, reducing their efficiency.

Chapter 2

Natural Gas Prices, Electric Generation Investment, and Greenhouse Gas Emissions

2.1 Introduction

Natural gas prices have fallen by over 65% from their high in 2008. This decline has been driven primarily by the large-scale expansion of hydraulic fracturing (fracking) for natural gas, which has transformed the US natural gas sector. Prior to fracking's development in the mid-2000's, production of natural gas was declining, and projected to decline further. With very large reserves of shale gas that can be fracked, firms are restricted in the number of productive wells they can drill only by the capability of their drilling rigs. Consequently, fracking has greatly increased the amount of natural gas produced in the US and the share of total US natural gas production from shale gas. From 2007 to 2013, total US natural gas production increased from 24.7 trillion cubic feet (TCF) to 30.0 TCF and shale gas production more than quintupled from 2 TCF to 11.9 TCF.[1] This long-term shift in the natural gas sector is still ongoing.

The reduction in natural gas prices due to the increase in supply from fracking has led to a large increase in natural gas consumption in the electric sector. Electric utility consumption of natural gas has increased by 1.3 TCF between 2007 and 2013, or roughly 20 percent.[2][3] Increased natural gas consumption has come at the expense of dirtier coal, causing a decrease in carbon emissions.

[1] EIA, http://www.eia.gov/dnav/ng/ng_prod_sum_dcu_NUS_a.htm.

[2] 2012's consumption was 2.3 TCF than 2007's, but consumption fell in 2013 amid slightly higher prices.

[3] Other uses have seen smaller changes. Residential consumption has increased by about 0.2 TCF and industrial consumption has increased by about 0.75 TCF (EIA, http://www.eia.gov/dnav/ng/ng_cons_sum_dcu_nus_a.htm). Net exports have increased by 2.5 TCF, though they remain negative (EIA, http://www.eia.gov/dnav/ng/ng_sum_sndm_s1_m.htm). The United States has negative net exports because it imports more natural gas than it exports, even when accounting for increased natural gas supply due to fracking. Most imported gas is from Canada.

Total carbon emissions from electricity generation declined from 284.7 thousand tons/hour in 2008 to 253.0 thousand tons/hour in 2013. The decline of 31.7 thousand tons/hour, equal to 11.1% of the total, was due to a variety of factors – increased renewable generation, slightly decreased demand, lower gas prices, and increased natural gas generating capacity.

This paper investigates the effects of lower gas prices, caused by increased natural gas production, on electric sector greenhouse gas emissions. I study the two primary channels that have been affected. I first examine how falling natural gas prices affect regional carbon emissions from electricity generation over the short run. Cheaper natural gas replaces relatively more expensive coal in the generation order; carbon emissions decrease because natural gas releases roughly half of the carbon that coal releases. I next consider the effect of falling natural gas prices on carbon emissions through unanticipated generation capacity additions.[4] Construction of gas-fired power plants has greatly exceeded projections made prior to the dramatic decrease in the natural gas price. Many of these new gas-fired power plants would not have been constructed if gas prices had remained high.

I empirically estimate the relationship between carbon emissions and natural gas prices using a flexible model that also includes electricity demand and a rich set of controls. My specification allows me to separately identify the effects of lower gas prices and increased gas-fired generation capacity. I also examine the interaction of these two effects. While carbon emissions reductions due to low gas prices are only available if prices remain low, reductions due to new capital stock will likely persist at moderately higher gas prices. This type of effect has been demonstrated before by Davis and Kilian (2011) in the home-heating market.

I exploit short-term variation in gas prices to identify the effect of gas prices on carbon emissions. There are two primary sources of this variation. The first is weather shocks, which influence gas prices in the short term. Unexpectedly cold weather forecasts boost demand for natural gas because many US homes are heated using gas. Unexpectedly temperate weather forecasts decreases demand. Finally, unexpectedly hot weather forecasts indicate increased air conditioning usage, increasing demand for electricity (and consequently increasing demand for gas). The second source of variation is production and storage reports. For example, unexpectedly high storage withdrawal, unexpectedly low storage injection, or unexpectedly low production reports will all

[4]During the previous period of low natural gas prices in the early 2000's there also was a large natural gas-fired capacity expansion.

increase the price of natural gas.[5][6]

I carefully control for the endogeneity of the price of natural gas. This endogeneity concern arises due to correlated demand shocks that may directly change both gas prices and electricity demand. For example, unseasonably warm winter weather may decrease both gas prices and electricity demand. This would cause my estimates to overstate emissions decreases attributable to lower gas prices. I control for this by including electricity demand directly in my specification. Additionally, it is possible that as gas prices fall, electricity prices may also fall (increasing the demand for electricity and, therefore, carbon emissions). Including electricity demand in my specification allows me to shut down this channel.

Next, I construct a counterfactual of what emissions would have been had natural gas prices remained at their higher levels prior to the large-scale application of fracking. In doing so, I control for renewable production and electricity demand levels. This allows me to answer my first research question; I find that lower gas prices caused 2013 carbon emissions to decrease by 14,700 tons/hour.

Falling natural gas prices may also influence carbon emissions through additions of new capital stock. This new capital stock may displace dirtier coal-fired power plants. I first determine the portion of new capital stock constructed in response to lower gas prices. This is a difficult question. I take three different approaches and conclude that 65-85% can be attributed to falling natural gas prices. In my first approach, I regress construction starts on gas prices and electricity demand growth. My second approach compares projections of capital additions from the EIA's Annual Energy Outlook with actual construction. This model makes its forecasts using aggregate data – it is a "macro" model. My final approach compares projections of capital additions from Form EIA-860 with actual construction. The form uses micro data submitted by utilities and independent power producers.[7] I use the range produced by these three approaches to estimate the amount of new capital stock constructed because of low gas prices. I also consider other potential causes of above-expectation gas-fired capacity construction. It is difficult to conclusively rule them out, but on balance the evidence points to lower gas prices caused by fracking.

To determine how newly constructed capacity has altered carbon emissions, I rely on the rela-

[5] Although gas prices have decreased over the long term as fracking has become more prevalent, variation caused by long-run supply changes is difficult to isolate from other trends that change emissions, such as macroeconomic conditions, increasing attention to energy efficiency, and technological improvements. For this reason, I use short-run variation.

[6] Additional details about the sources of gas price variation are available in Appendix B.1.

[7] The Annual Energy Outlook is based, in part, on data from the Form EIA-860. It is combined with other data in order to make projections.

tionship between carbon emissions and electricity demand. Identification of this relationship also relies on short-term variation, and weather again is a key source of this variation.

In order to construct counterfactual emissions where new gas-fired capacity does not exist, I first determine hour-by-hour electricity generation and carbon emissions from the new capital stock. Then, I use the coefficients on electricity demand to determine what marginal emissions would have been if this electricity was instead generated by the existing power plant fleet. The difference between actual emissions and counterfactual emissions reveals the emissions savings caused by the new capital stock. I estimate 2013 carbon reductions from new capacity to be an additional 2,100 tons/hour.

There are several related studies of the relationship between gas prices and carbon emissions. Lu, Salovaara and McElroy (2012) examine the effect of gas prices on emissions by EPA region using monthly data over a two-year period. They find that between 2008 and 2009, carbon emissions fell by 8.76% in the electric sector, 4.3% of which was due to falling natural gas prices. This paper uses hourly data over a seven-year period to provide more robust and comprehensive estimates and also looks at new gas-fired construction.[8] Cullen and Mansur (2014) estimate the relationship between gas prices and carbon emissions in order to analyze the industry response to a tax on carbon emissions. While their analysis relates emissions to the gas price, they do not specifically address the effect of recent gas price declines on carbon emissions or the effects of new capital stock. [9,10]

My paper fits in a broader literature that estimates the effects of natural gas prices on the power sector (Holladay and LaRiviere, 2014; Linn, Muehlenbachs and Wang, 2014). It is also relevant within the literature that estimates greenhouse gas emissions from the electric sector (Kaffine, McBee and Lieskovsky, 2013; Callaway, Fowlie and McCormick, 2015; Linn, Mastrangelo and Burtraw, 2014; Cullen, 2013; Novan, 2014).

The paper proceeds as follows. I provide some institutional background to help frame the analysis (Section 2.2). Next, I analyze the amount of new generation capacity prompted by low gas prices (Section 2.3). I briefly discuss my data (Section 2.4). I detail my empirical model and

[8]Hourly data is less reflective of medium-run trends, such as climate policy, than more smoothed monthly data. Even with hourly data, there is still a concern that medium-run trends affect both the gas price and emissions; the added years of data allow for the inclusion of a time trend and monthly fixed effects to help control for this. Using NERC interconnections alleviates concerns about cross-region electricity trading.

[9]Lafrancois (2012) estimates the potential effect of gas-fired power plants constructed before 2006 on carbon emissions.

[10]Engineering models such as Venkatesh et al. (2012) have also looked at this topic. They use simplified dispatch models to estimate how the marginal supply curve will change. While this approach has its advantages, it may be less precise for marginal changes. Transmission losses and costs, bottlenecks, ramping costs, market power, and outages may make it so that some power plants are more likely to provide power than their marginal cost would suggest.

results in Sections 2.5 and 2.6. I consider alternative specifications to check the robustness of my results, then discuss, and conclude (Sections 2.7, 2.8, and 2.9).

2.2 Background

There are a several institutional details that are important to my analysis. For the vast majority of American consumers, electricity prices do not vary in real time. Thus, electricity demand does not adjust in real-time in response to changing wholesale electricity prices – it is almost completely inelastic in the short-run. Over the medium-run, demand has the potential to adjust in response. Electricity prices have been relatively stable in real terms recently. Between 2007 and 2013 the real annual national average price of electricity fluctuated between 9.35 and 9.98 cents per kilowatt-hour (EIA).[11] Electricity consumption has also been relatively constant. In the discussion section, I examine the potential medium-run demand response.

It is important to note that manipulation of the price of natural gas is not a concern for my estimation. Gas power plants are price-takers and are unable to manipulate the price of natural gas. The natural gas market is large and groups of power plants are unable to substantially move the market. Additionally, it is the case that power plants with long-term contracts are not required to burn gas at any individual time. Thus, whether a plant has a favorable or unfavorable long-term contract (or no long-term contract) has little bearing on whether they decide to supply electricity. The opportunity cost of producing electricity is the spot price of natural gas that firms pay.[12]

2.2.1 Interconnection Analysis

For my analysis, I focus on the NERC interconnection level as in Graff Zivin, Kotchen and Mansur (2014). Figure 2.1 illustrates the location and boundaries of the three interconnections and regions within each that the NERC oversees. The interconnections are largely separate entities, with minimal electricity trading between each interconnection. The Western interconnection (WECC) covers most of the territory from New Mexico up to Montana and west. The Texas interconnection (TRE) covers most of Texas. The Eastern interconnection is subdivided into six different regional entities that comprise the rest of the United States.

Figure 2.2 shows power flows between different regions. For reference, one million megawatt hours over the course of a year are equivalent to an average of roughly 115 megawatts during every hour. Electricity flows between interconnections (circled) are very small. Regions with

[11] This is across all segments, not just residential consumption. Monthly data has a little more variation. I use 2010 price levels.

[12] For more on gas markets, please see EIA (2001).

Figure 2.1: NERC Interconnections

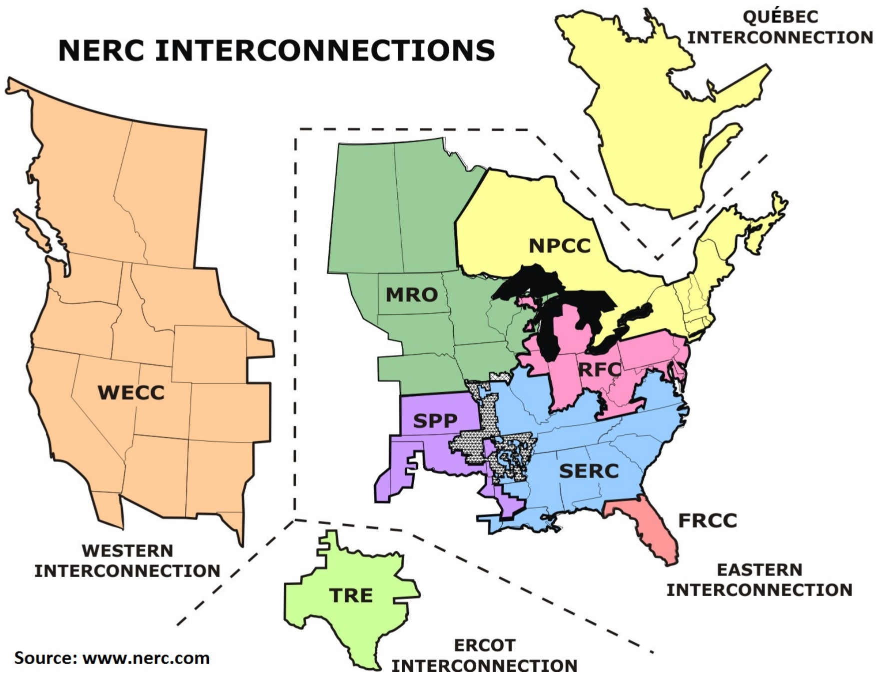
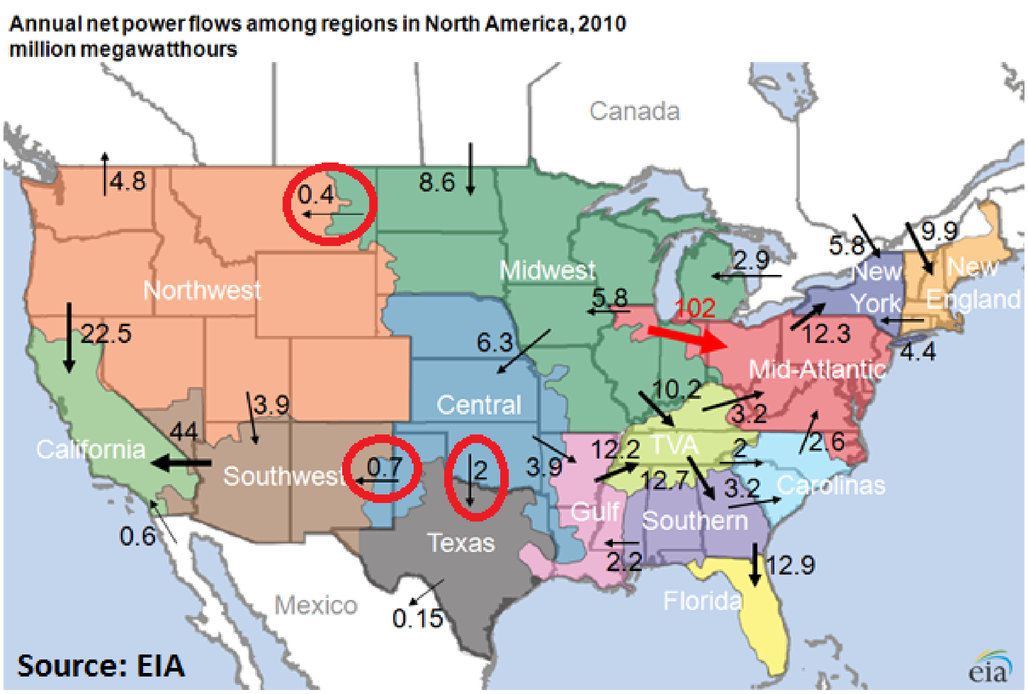


Figure 2.2: Regional Power Flows



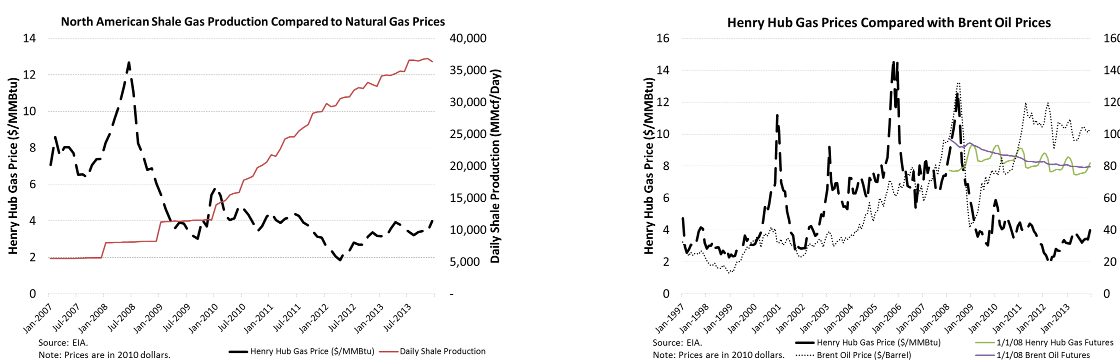
large amounts of power trading, like the Chicago area with eastern parts of the RFC (denoted by the 102 million megawatt hours), would be inappropriate fits for my model. While Canada does trade power with the United States, the lines are primarily transmitting hydroelectrically generated electricity and are full at most hours. Thus, it should have limited effect on the analysis.

A large amount of regional trading threatens clean identification of this relationship. To understand this, consider two regions, the Midwest Reliability Organization (MRO) and the Southwest Power Pool (SPP). MRO has large amounts of coal capacity, while SPP has a mixture of coal and natural gas. Assume that the regions trade freely and have large amounts of transmission capacity between them. Additionally, assume electricity demand remains fixed. As the price of natural gas decreases, more gas and less coal will be burnt. This would mean that in aggregate carbon dioxide emissions would decrease. However, it could be the case that power generation has increased in SPP and decreased in MRO, with SPP sending excess generation to MRO. SPP would then show an increase in emissions at lower gas prices, while MRO shows a larger than deserved decrease. A system with minimal trading prevents this potential identification issue.

2.2.2 Gas Prices and the Dispatch Curve

There has been substantial variation in the natural gas price over the last several years. The spot price of gas at Henry Hub, the most important trading location, does an excellent job of capturing this variation. United States gas markets are fairly integrated, with most other locations trading at a basis against Henry Hub. For example, natural gas in Chicago is generally about 10 cents per MMBtu (2-5%) more expensive than natural gas at Henry Hub.

Figure 2.3: Gas Prices, Oil Prices, and Shale Production



(a) (b)

Figure 2.3 examines changes in the natural gas market that have been occurring since 2007. The panel on the left (a) plots the relationship between the Henry Hub Natural Gas Spot Price and the quantity of shale gas that is produced in North America. It demonstrates that as the supply of shale gas has increased, the price of natural gas has decreased. The vast majority of shale gas is drilled using hydraulic fracturing. Note that the first three years of shale data were only collected annually, though quantities are generally small. Additionally, the gas price is a monthly average. This figure depicts the long-term trend, though it obscures the day-to-day variation in gas prices that is key to my identification strategy.

The panel on the right (b) plots the price of gas at Henry Hub against the Brent oil price.[13] Oil and gas are energy sources that are, to a certain extent, substitutable. Prior to the large-scale implementation of fracking, oil and gas closely tracked each other, with a barrel of oil being about ten times as expensive as an MMBtu of natural gas[14]. The graph starts in 1997 when the EIA Henry Hub spot price time series begins, though the relationship in the early 90's (using a different measure of the natural gas price) was also strong. In 2008 there was a recession-induced decline in the prices of both fuels. Shale gas production greatly increased during the recovery, and the relationship between gas and oil prices fractured. Oil prices surged back to pre-recession levels, while gas prices continued their decline. By the end of 2013, a barrel of oil was now *twenty-five* times as expensive as one MMBtu of natural gas.[15]

If macroeconomic conditions were the only important changes in energy markets, gas prices would likely have rebounded similar to the oil price rebound (Hausman and Kellogg, 2015). The fractured relationship is possible because oil is traded on a global market, whereas natural gas markets are regional (Kilian, 2015). Fracking has also produced a US oil boom, but it hasn't had as large of an effect on the world price of oil because the global oil market is very large and the oil boom began later. Excess natural gas within the US is unable to be exported in large quantities outside of North America because of a lack of infrastructure and high transportation costs. Instead, it is consumed locally at much cheaper prices.

Also included in panel (b) are futures curves from January 2008 for natural gas and crude oil. The futures curves show that financial markets expected gas and oil prices to retain their historical relationship over the 2009-2013 period. Financial markets also did not anticipate the large decline in gas prices.

As discussed in the introduction, total gas production in the US increased by 5 trillion cubic feet per year, or 20%, between 2007 and 2012. The combination of a large increase in quantity

[13] Brent oil is the major world oil price. Oil prices in Cushing, Oklahoma are similar, though they have been slightly lower because of pipeline constraints.

[14] That is, oil was about twice as expensive on a per-MMBtu basis

[15] For more on the relationship between oil and gas prices, please see Villar and Joutz (2006); Ramberg and Parsons (2012) and EIA (2012).

supplied and a large decrease in the price of natural gas is strongly suggestive of a large rightward shift of the natural gas supply curve.

The price of natural gas is an important factor in electric-sector carbon emissions. In some areas of the country, firms bid in real-time to determine who is going to supply the marginal kilowatt-hour. Nuclear and renewable power plants have very low marginal costs. As a result, nearly all marginal power is provided by either coal or natural gas-fired plants. In other areas of the country, a centralized dispatch authority determines which plants produce power. One of the authority's main objectives is to minimize generation costs. Changing fuel costs will prompt a dispatch authority to adjust the generation mix.

Fuel is the primary variable cost at fossil fuel power plants. Moderate to high natural gas prices usually cause coal to have lower variable costs than coal, making it the first fuel called upon to generate electricity. At lower gas prices, the marginal cost of electricity generated from gas will decrease and gas will begin to displace coal in the generation order. This switching between coal and natural gas is the key mechanism driving the results in this paper. Switching is able to happen within a period of hours.

2.3 Cheap Gas and Gas-Fired Capacity Additions

Most forecasters and industry analysts were expecting only very minor gas-fired capacity increases between 2010 and 2013. Instead, 25.9 GW of gas-fired capacity was added over this timeframe. It is difficult to isolate the precise effect of fracking on new gas-fired capacity additions. I take three approaches to answering this question, which all yield similar results. First, I run a set of simple regressions that estimate the relationship between the gas price and construction starts. One weakness of this approach is the limited sample. Next, I examine projections made by the EIA in their Annual Energy Outlook projections. Finally, I consider data about potential projects that are filed with the EIA using their EIA-860 form.[16] I look at these projections because they were made before fracking; differences from the projections can plausibly be ascribed to fracking. I estimate that roughly 65%-85% of these additions likely would not have happened if the gas price had remained at 2008 levels.

A gas-fired power plant takes between 18 and 36 months to construct. Natural gas prices crashed in mid-2008, suggesting that the earliest gas plants built because of low gas prices would likely have come online in 2010. While it is likely that a few suspended projects were restarted and completed by 2009, I do not consider these plants.

[16]Note that the AEO projections are based, in part, on the raw EIA-860 data.

My analysis includes both combined cycle and conventional combustion turbines.[17] New combined cycle plants primarily supplant less efficient coal-fired plants, while new combustion turbine plants could displace the least efficient coal-fired plants during shoulder periods, as well as less efficient (oil or gas) peaker plants during peak hours. However, emissions savings from new combustion turbines are likely to be minimal – new combined cycle plants likely drive the results in this paper.

It is difficult to disentangle the effects of natural gas prices from contemporaneous trends such as state renewable portfolio standards, changing environmental regulations, or the great recession. I briefly consider the effect each of these trends might have had on gas-fired generation construction. While not definitive, these considerations support the theory that gas prices were the major driver of gas-fired capacity additions.

2.3.1 Construction Starts Regression Analysis

I first consider a regression-based approach to determine the relationship between gas prices and estimated gas-fired construction starts. Prior to the regressions, I use data from Form EIA-860 to estimate construction starts.[18] The data summarizes construction completions (e.g., 20.1 GW in 2004 and 14.8 GW in 2005); I use an 18-month lead to estimate construction starts for each year (e.g. 17.4 GW in 2003).[19] Using lead completions data as an estimate of construction starts instead of actual construction starts allows me to capture some intermediate effects. For example, some power plants are begun, but later "indefinitely postponed." Lead completions estimates appropriately account for plants that were indefinitely postponed and later restarted, as well as the lower likelihood of such postponements when firms expect a long-term supply of inexpensive natural gas.

The Annual Energy Outlook also reports construction completions. Their data is based on the EIA-860, though it is compiled differently.[20] As a check, I use both the EIA-860 and AEO datasets in my analysis.

Using data from 2000-2011, I estimate the annual relationship between construction starts (in

[17] Approximately 65% of this new gas-fired capacity was from combined cycle plants.

[18] The raw EIA-860 data are aggregated in the Electric Power Annual (EPA). I use aggregated data in the EPA because disaggregated data on gas-fired construction starts from the EIA-860 are unavailable for some years due to changes to EIA's data collection procedures.

[19] I choose an 18-month lead because it allows for the best fit with the available micro data on construction starts. In Appendix B.2 I provide a range of alternative lead times; results are similar.

[20] The raw EIA-860 data are aggregated in the Electric Power Annual. I use aggregated data because disaggregated data on gas-fired construction starts from the EIA-860 are unavailable for some years due to changes to EIA's data collection procedures.

megawatts) and electricity demand growth and the price of natural gas. Specifically, I estimate:

$$LoggedConstructionStarts(C_t) = \alpha_0 + \beta_1 P_t^{NG} + \beta_2 ElecGrowth_t + \epsilon_t \tag{2.1}$$

The price of natural gas and electricity demand growth are the two primary drivers of gas-fired capacity investments. Gas prices determine the marginal cost of operating plants, while demand growth helps determine future wholesale electricity prices. I consider two variations of the dependent variable, using either AEO or EIA-860 data.

A limitation of this estimation is that the number of data points in this time series is only 11 or 12, depending on the data source.[21] More reliable estimates would result from including additional controls, making the independent variables more flexible, and adjusting the standard errors. Data limitations prevent these adjustments. Nevertheless, this estimation allows for a rough look at the relationship between gas prices and construction starts. I have summarized the results in Table 2.1.

While the magnitude of the gas price coefficient varies across the regressions (including the ones in Appendix B.2), it is consistently negative. To determine the counterfactual construction, I first determine the difference between actual gas prices in each year and the counterfactual (no-fracking) gas price from 2008. I then adjust construction starts in each year down by the counterfactual gas price difference multiplied by the gas price coefficient. This allows me to determine, in rows [c] through [g], what counterfactual construction would have been.

In 2009, construction estimates do not change because of the way this analysis is constructed – it takes more than 12 months to have an effect. However, starting in 2010, counterfactual construction is frequently lower than it otherwise would have been. Much of the time it is close to zero or negative. I interpret negative construction to mean that it is very undesirable to build a gas plant, not that gas plants are being decommissioned. Years in which there is positive construction are highlighted.

Finally, I compare actual plant construction from 2010 to 2013 with counterfactual plant construction. I take the 25.9 GW of stock that was actually constructed and subtract construction in any year with a positive counterfactual. For example, in column [2], I subtract (1.5 + 3.6 + 2.8 + 1.7 =) 9.6 GW that would have been built even if gas prices remained high. Negative construction is treated as a zero.

Depending on the specification, this analysis suggests either 16.2 or 22.1 GW of gas-fired capacity was constructed that would not otherwise have been built. Note that the EIA-860 plant

[21] In Appendix B.2 I include a scatterplot of these points, as well as a line of best fit.

Table 2.1: Regression Results: Counterfactual Plant Construction

Estimated Counterfactual Plant Construction
Results from Regression Analysis Using an 18 Month Lag

Item		Using AEO Plant Data	Using EIA-860 Plant Data
		[1]	[2]
Gas Price Coefficient	[a]	-0.13	-0.21
Standard Error	[b]	0.12	0.12

Counterfactual Plant Construction (Gigawatts)

Year		Using AEO Plant Data	Using EIA-860 Plant Data
2009	[c]	10.9	9.4
2010	[d]	1.5	3.2
2011	[e]	3.6	0.1
2012	[f]	2.8	< 0
2013	[g]	1.7	
Actual Construction (2010 - 2012/2013)	[h]	25.9	25.4
Construction Caused by Fracking	[i]	**16.2**	**22.1**

Notes:
[b]: The standard error does not address autocorrelation.
[g]: 2013 EIA-860 data is currently unavailable. Placeholder data suggests counterfactual construction is near zero and actual construction is greater than zero. This suggests I am underestimating the effect of low gas prices.
[i]: = [h] - Sum([d] to [g]) if positive & shaded. That is, [i] is construction that would not have occurred with higher gas prices.
Sources: EIA-860 data summarized in the Electric Power Annual and Annual Energy Outlook Data.

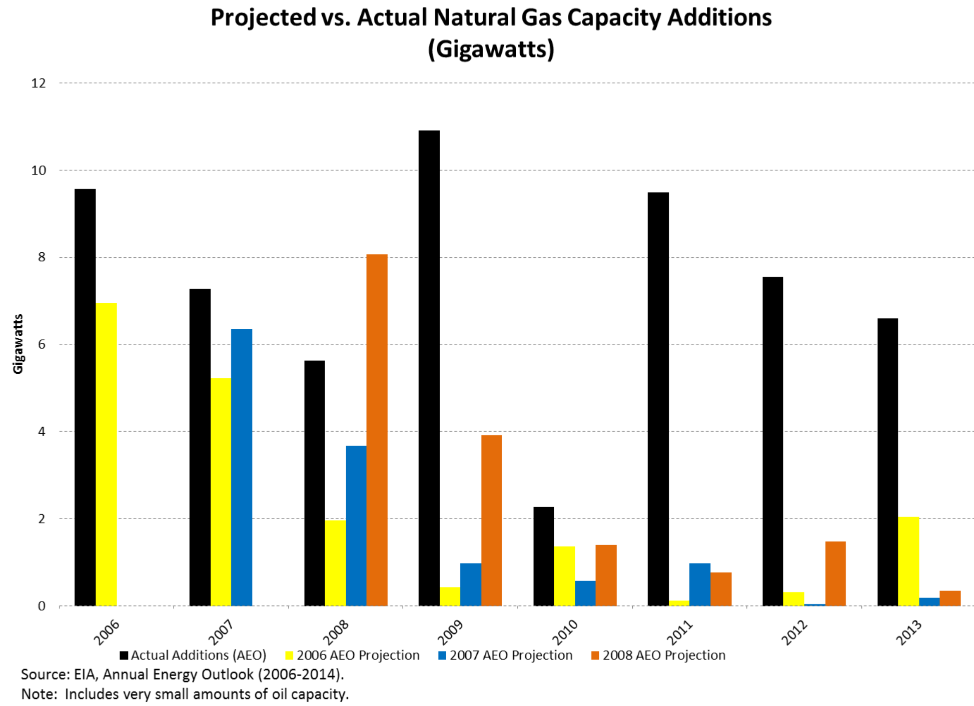
data do not include 2013 construction or counterfactual construction.

2.3.2 Annual Energy Outlook Projections

In the mid-2000s, the EIA estimated that there would only be very modest investment in natural gas-fired electric generation capacity. The available capital stock would be mostly sufficient to meet growth in electricity demand. Further, it would not be profitable to invest in new capacity while old capacity was working well. The 2007 Annual Energy Outlook (AEO) projections suggested that there would be roughly 2 GW of natural gas capacity added during the 2010-2013 timeframe. As Figure 2.4 shows, this was not a one-year aberration; projections in surrounding years were also very similar.

However, the solid black column in Figure 2.4 indicates that actual capacity additions were substantially above initial projections. The 25.9 GW of capacity that was built between 2010 and

Figure 2.4: Projected vs. Actual Natural Gas Capacity Additions



2013 is much higher than these projections.

I next look to see how close previous AEO projections were to actual construction. It is possible that the AEO's black box model regularly underestimates short to medium-run gas-capacity additions. To be conservative and allow for this possibility, I adjust the 2006-2008 AEO projections up by the amount that previous projections missed by. I view this as conservative in part because AEO projections are intended to be unbiased. Table 2.2 details this analysis.

I look at both five-year (columns [1] to [3]) and ten-year projections (columns [4] to [6]) in the same way. In the top panel I analyze projections made between 2001 and 2004 to determine how accurate they were. These projections are mostly before the advent of fracking and are mostly free from its influence. In the five-year projection, total construction averaged 147% above projection. Ten-year construction averaged 22% above projection.[22]

In the bottom panel, I analyze 2006-2008 projections for construction between 2010 and 2013. All three years projected minimal construction during these years. To adjust for previous under-projections, I multiply the 2006-2008 projections by ratio of actual/projected construction that I

[22]Note that these results are driven in part by 2004 projections (row [d]), which are most likely to be influenced by fracking. I view the inclusion of 2004 projections as conservative.

Table 2.2: AEO Analysis: Counterfactual Plant Construction

Estimated Plant Construction Due to Fracking
Differences from AEO Projections
(Gigawatts)

Projection Year		5-Year Projection	5-Year Actual Construction	Percent Constructed Above (Below) Projection	Up to 10-Year Projection	Up to 10-Year Actual Construction	Percent Constructed Above (Below) Projection
		[1]	[2]	[3]	[4]	[5]	[6]
Pre-Fracking Projections					**Mostly Pre-Fracking Projections**		
AEO 2001	[a]	78.3	149.9	92%	189.7	185.6	-2%
AEO 2002	[b]	91.2	151.2	66%	195.1	186.8	-4%
AEO 2003	[c]	49.9	97.9	96%	118.9	133.8	12%
AEO 2004	[d]	13.8	60.4	336%	53.0	97.2	83%
Pre-Fracking Average	[e]	58	115	**147%**	139	151	**22%**

Amount of 2010-2013 Construction (26 GW Total) Attributable to Fracking

		2010-2013 Projection	Projections Adjusted for Errors	Gigawatts Above Projection	2010-2013 Projection	Projections Adjusted for Errors	Gigawatts Above Projection
AEO 2006	[f]	3.9	9.5	16.4	3.9	4.7	21.2
AEO 2007	[g]	1.8	4.4	21.5	1.8	2.2	23.7
AEO 2008	[h]	4.0	9.9	16.0	4.0	4.9	21.0
Average	[i]			**18.0**			**22.0**

The upper panel determines how accurate previous AEO estimates of natural gas plant construction were. This information is used in the lower panel to adjust the projections for 2010-2013 construction. Because previous projections overestimated construction, these estimates are conservative relative to the baselines outlined in the 2006-2008 AEOs.
The "Projections Adjusted for Errors" columns ([2] & [5] on the lower panel) are constructed by multiplying the actual projections (in columns [1] & [4]) by the percent that previous projections were overestimated (from the upper panel).
[i]: Depending on whether I use a 5-Year or 10-Year error adjustment, I estimate between 18 and 22 GW of capacity was constructed that would not have been if gas prices were higher. See text for additional details.
Source: Annual Energy Outlook.

calculate in the upper panel (row [e], in bold). Even after I adjust for previous errors (columns [2] and [5]), the expected level of construction was much lower than actual construction. Adjusting for previous projection errors, this analysis suggests between 18 and 22 gigawatts of natural gas-fired capacity was constructed that would not otherwise have been.

2.3.3 Raw EIA-860 Projections

Proposed electricity plants are required to file the EIA-860 form if "[t]he plant will be primarily fueled by energy sources other than coal or nuclear energy and is expected to begin commercial operation within 5 years." This form details proposed plants, which are in various stages of planning or construction, but are not generally certain to be completed. When using the raw EIA-860 data, I view it as a soft cap on the possible number of projects that will be built over the next 5 years.[23] To construct a plant within the next five years that is not already in the database, a firm would need to complete the siting, planning and construction phases. This can be done, but it requires a very smooth process.

To determine how many of these projects are completed during pre-fracking (normal) times, I look at summaries of EIA-860 data from 2001 through 2003. As Table 2.3 shows, on average 59% of potential projects were completed during these years. In contrast, when I look at projections from 2006 through 2008, years that mostly overlap with fracking, I see that 101% of potential projects are completed. That is, during regular times, half of all projects are completed. When gas-fired plants become much more profitable because their marginal costs tremendously decline, slightly more projects are completed than were planned.

This suggests that without fracking (and lower gas prices), roughly one half of all projects would have been completed – the other half was induced by very cheap natural gas prices. Looking at the most recent projection for which I have five years of actual construction data (2007), I estimate that 17.5 GW of additional capacity were induced by cheap natural gas.

2.3.4 Alternative Explanations

There are a number of possible alternative explanations for the surge in gas-fired construction. I now briefly consider several of them. While I am unable to conclusively rule them out, they do not appear to be the main driver of new plant construction.

[23] This raw data is summarized in the Electric Power Annual.

Table 2.3: EIA-860 Analysis: Counterfactual Plant Construction

Estimated Counterfactual Plant Construction
EIA-860 Registered Projects
(Gigawatts)

Projection Year		5-Year Potential Projects	Actual Construction	Number Not Constructed	Percent Constructed
		[1]	[2]	[3]	[4]
Potential Projects are Pre-Fracking					
2001	[a]	285	146	139	51%
2002	[b]	160	96	63	60%
2003	[c]	88	58	30	66%
Potential Projects are Mostly During Fracking					
2006	[d]	40	40	(0)	101%
2007	[e]	41	42	(1)	103%
2008 (Four Years)	[f]	35	35	(0)	100%
Pre-Fracking Average	[g]	178	100	77	59%
Fracking Average	[h]	39	39	(1)	101%
Increase in Completed Projects	[i]				**42%**
Estimated Gigawatts Attributable to Fracking	[j]	**17.5**			

[g]: = Sum([c] to [d]).
[1,j]: To estimate GW attributable to fracking, I compare the percentage completed during fracking (101%) with the percentage completed during normal times (59%). I attribute the 42% difference to cheaper gas prices. 2007's 5-year projection was that 41 GW were potentially under construction, 42% of this is 17.5 GW. Note that this does not account for delayed retirements or unexpected capacity increases.
Sources: EIA-860 data summarized in the Electric Power Annual and Annual Energy Outlook Data.

State Renewable Portfolio Standards

Renewable Portfolio Standards have been enacted at the state level in twenty-nine of the lower 48 states (and the District of Columbia). Nineteen states, generally in the mountain region or southeast, had either voluntary goals or no legislation.[24] While standards vary across states, they broadly seek to increase the amount of power generated from renewable sources. These standards likely disincentivize gas-fired construction because renewable generation will cover much of future electricity growth. However, when building new fossil-fuel plants, they could also incentivize additional gas-fired power plants (at the expense of coal-fired plants) because gas-fired

[24] http://www.eia.gov/todayinenergy/detail.cfm?id=4850.

generation better complements the less predictable nature of renewable power. Using Texas data, Dorsey-Palmateer (2014) finds that the primary effect of wind generation is to reduce fossil fuel consumption. That is, the first effect would likely outweigh the second, and in the absence of renewable portfolio standards, gas-fired construction would likely have been even larger.

I now look to see if a disproportionate share of construction was in states with renewable portfolio standards. The twenty-nine states with standards contained 72% of the US population. They also constructed 76% of the 227 new gas-fired units built over the 2010-2013 period (note that plants can have multiple units). This (very broad) overview does not suggest a large effect due to renewable portfolio standards, as there was also substantial construction in states without these standards.

Great Recession

The Great Recession began in December of 2007 and ended in June of 2009. In this section, I have used projections that were issued between 2005 and early 2008. For example, section 2.3.2 uses the 2007 AEO projection issued in early 2007 before the onset of the recession. Following the onset of the recession, capital expenditures across the US economy fell substantially. If the construction projections accounted for the upcoming Great Recession, they likely would have predicted an even lower amount of new gas-fired power plant construction. That is to say, while the effects of the Great Recession are not captured in this analysis, the recession likely muted the effect of lower gas prices.

Changing Environmental Regulations

The Regional Greenhouse Gas Initiative (RGGI) involves ten states in the Northeastern US (plus PA as an observer). It went into effect in 2009 in an attempt to limit carbon emissions. Similarly, California implemented a cap & trade program in 2013. These programs will increase the cost of emitting carbon (up from zero) and some of these costs may be passed on to consumers in the form of higher electricity prices. The effect on natural gas-fired plants is ambiguous – they are cleaner than coal-fired generation, but dirtier than renewables. The twelve states (Northeast plus California) make up about 33% of US population, and have built 38% of new gas-fired generation units. I do not interpret this finding as evidence that carbon regulations have been driving gas-fired investments.[25]

[25] Interestingly, California has built more than its share of new generation while the Northeast has built less. This could be related to population growth that is above the national average in California and well below the national average in the northeast.

Coal-fired power plants are the largest source of mercury emissions in the United States. It is possible that changing mercury regulations have influenced the decision to build gas-fired plants. At the national level, Mercury and Air Toxics Standards (MATS) are being developed by the EPA. They were originally proposed in March 2011, but have been under revision since then. As of April 2015, the standards look like they will be upheld.[26] It is possible that these standards influenced borderline plants to continue completion. It is unlikely that MATS caused new plants to be conceived and constructed between 2010 and 2013 because of the uncertainty surrounding the revision and the amount of time required to build a new power plant. Additionally, 2011's large amount of construction completions, which was likely unaffected by the MATS proposal, provides evidence that new construction was economic without the benefit of MATS. However, I cannot conclusively rule out MATS as a driver of gas-fired construction.

The Clean Air Interstate Rule (CAIR) was originally proposed by the EPA in 2003 with the aim of reducing emissions of particulate matter, nitrogen oxide, and sulfur dioxide. After a lengthy legal battle, CAIR was remanded in 2008 and the EPA was ordered to address several problems with the regulation. The EPA finalized the Cross-State Air Pollution Rule (CSAPR) in 2011, and phase I took effect at the start of 2015. The primary effect of the law is to reduce pollution from coal-fired power plants through additional technological controls or reduced generation. Natural gas plants emit less conventional (non-carbon) pollution when compared with coal-fired plants. As a result, I expect the net effect of the regulation to promote gas-fired and renewable generation at the expense of coal. CSAPR and/or CAIR were designed to affect 31 states and the District of Columbia. These states comprised 75% of the US population, but only 60% of the new gas-fired construction. I interpret this as evidence that CSAPR is not the primary driver of new gas-fired construction. States where CAIR/CSAPR promote gas-fired generation had a *lower* than representative percentage of new gas-fired construction.

Alternative Explanations Review

There have been several important changes to the electricity sector over the previous fifteen years. State renewable portfolio standards and the Great Recession likely disincentivized gas-fired construction. Changing environmental regulations may have promoted gas-fired construction. However, the affected states do not have a disproportionate share of construction. The regulations also do not appear timed such that they would substantially affect construction projections from 2007.

[26] http://green.blogs.nytimes.com/2011/12/21/e-p-a-announces-mercury-limits/, http://www.epa.gov/mats/actions.html.

2.3.5 Inference

Estimating the amount of gas-fired construction that is due to lower gas prices is a difficult problem. In this section, I have taken three approaches. Using a construction regressions approach, I estimate that between 16.2 and 22.1 GW of gas-fired capacity were added because of low gas prices. Using differences from AEO projections, I estimate total additions of 18.0 to 22.0 GW. Finally, using the raw EIA-860 data, I estimate additions to be 17.5 GW. All three approaches produce similar estimates; I estimate that between 65% and 85% of total additions were prompted by low gas prices. These estimates are also consistent with the intuition that greatly reducing marginal production costs (through lower gas prices) will incentivize firms to increase production capacity. However, it is difficult to disentangle the effect of natural gas prices from other factors affecting these large capital expenditures. The remainder of this paper turns to studying changes in carbon emissions from both existing and newly-constructed plants.

2.4 Data

Emissions data are collected by the EPA using the Continuous Emissions Monitoring System (CEMS). CEMS collects emissions data from all fossil fuel power plant units that have generation capacity of 25 megawatts or greater. Most power plants generate several hundred megawatts. Only very small generators (producing small amounts of pollution) are not included; CEMS covers the vast majority of pollutant-emitting electricity generation in the United States.[27] I use hourly data over the 2007 to 2013 period. Figure 2.5 summarizes carbon dioxide emissions from these power plants and illustrates the seasonality of electricity generation. Carbon emissions decrease by a little more than 10% over the time period. A graph of fossil fuel electricity generation by interconnection looks similar, though electricity generation has remained relatively constant.[28]

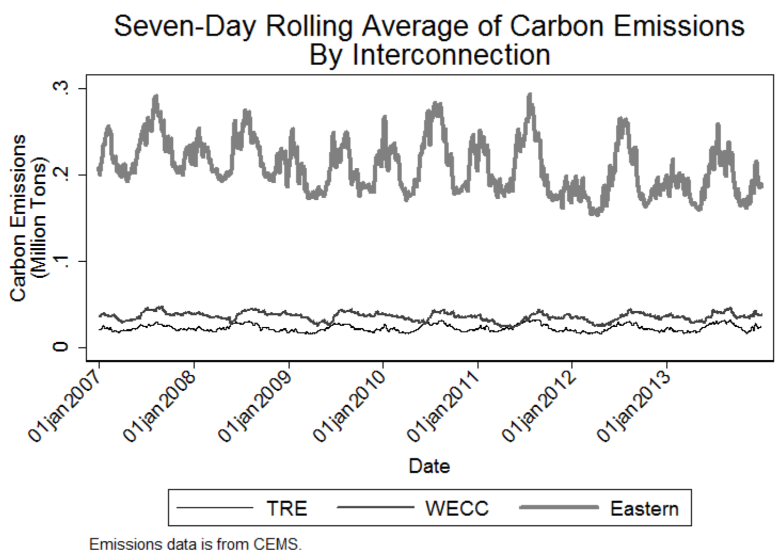
I use hourly electricity demand data from FERC Form 714. Planning areas are geographic zones that coordinate electricity load to meet demand. The FERC requires each planning area to submit this report annually. I map planning areas to NERC interconnections and then aggregate the data by interconnection, allowing me to control for changing demand.[29]

[27] I use generators labeled "Electric Utility," "Cogeneration," "Small Power Producer" or "Institutional." I consider this the backbone of the electric grid. I exclude a range of industrial plants like "Pulp & Paper Mill" or "Cement Plant" as they frequently do not list electricity generation, but do emit pollutants.

[28] This is anomalous. For several decades prior to the time period, electricity demand grew fairly steadily by a couple percent every year.

[29] For regions where independent system operators (ISOs) report separately from utilities, I only include data from the ISOs. This prevents double counting. For example, this means that the northeast is comprised only of data reported by the NYISO and NEISO, California's data is predominantly from CAISO, etc.

Figure 2.5: Seven-Day Rolling Average of Carbon Emissions



The EIA requires electricity generators to report monthly information via EIA form 923. I aggregate and use monthly net generation from renewable power plants.[30] Because renewable generation does not emit carbon dioxide or sulfur dioxide, it is not captured by CEMS.

I include data from the National Weather Service on heating degree days (HDD) and cooling degree days (CDD).[31] I take population-weighted averages for each interconnection.

Finally, I use natural gas spot price data that are collected by the EIA through Thomson Reuters. They track the natural gas price at Henry Hub.

2.5 Empirics

My analysis is essentially estimating a production function for carbon emissions. I have panel data on the electricity-generation industry over a period of seven years and am able to repeatedly

[30]Specifically, I use fuel codes for nuclear, hydroelectric, solar, geothermal, and wind power. This is consistent with Cullen and Mansur (2014). Less than 1% of generation is reported as a "State-Fuel Level Increment" without a NERC region. I assign this data to NERC regions. Results are similar if it is omitted.

[31]The National Weather Service defines HDD and CDD: "A mean daily temperature (average of the daily maximum and minimum temperatures) of 65F is the base for both heating and cooling degree day computations. Heating degree days are summations of negative differences between the mean daily temperature and the 65F base; cooling degree days are summations of positive differences from the same base. For example, cooling degree days for a station with daily mean temperatures during a seven-day period of 67, 65, 70, 74, 78, 65 and 68, are 2, 0, 5, 9, 13, 0, and 3, for a total for the week of 32 cooling degree days." http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/ddayexp.shtml.

view their emissions decisions. Given gas prices, electricity demand, and other control variables, I estimate the causal effect of gas price shocks on carbon emissions. I also estimate the causal effect of newly constructed gas-fired capacity on carbon emissions.

My identification assumption is that short-run changes in gas prices are uncorrelated with carbon emissions except through dispatch changes. After including appropriate controls, these price changes are orthogonal to other determinants of carbon emissions. Similarly, when estimating the causal effect of newly constructed gas-fired capacity, I assume that the electricity demand coefficients represent the marginal emissions from the power plants they are replacing. Gas price and electricity demand changes are exogenously caused and the resulting errors are uncorrelated with carbon emissions.

I run my analysis separately for each hour of the day. As Graff Zivin, Kotchen and Mansur (2014) show, marginal emissions can vary widely from hour to hour. If new gas-fired generators are running overnight they will likely be providing baseload power and replacing coal power plants. This will have a large effect on emissions. However, if the new generators are primarily running during peak hours of demand, they could just be replacing older gas-fired plants, providing minimal emissions savings.

I focus on the interconnection level (Western, Eastern, & Texas). Electricity demand is reported at the planning area. Due to changes in planning area geography, some planning areas move from one region to another region or cover multiple regions during my time period. For example, MISO (a planning area) covers parts of MRO, RFC, and SERC. Because of this and because of substantial trading across regions (Section 2.2.1), I do not report individual regional estimates for the Eastern interconnection.

2.5.1 Decrease in Emissions from Switching

My primary specification is run at the daily level (t) and is estimated separately for each interconnection and hour.[32] It estimates the relationship between total (carbon) emissions (TE_t) aggregated across the interconnection and the national price of natural gas (P_t^{NG}), controlling for interconnection-level electricity demand (Q_t^E), renewable electricity generation ($Renewables_t$), Heating Degree Days (HDD_t), and Cooling Degree Days (CDD_t). I also include a flexible time trend ($Date_t$) and month of year fixed effects (D_m). Finally, I include an interaction term between the gas price and the demand spline. I use a cubic spline, s(), with six knot points to allow for

[32]That is, each of the three interconnections runs the specification twenty-four times, for a total of seventy-two regressions.

flexibility in the relationship between emissions and the price of natural gas, electricity demand, renewables, and the time trend. The shape of the spline does not strongly depend on the number of knots. I choose six to be consistent with Cullen & Mansur (2014).[33]

$$\begin{aligned} TotalEmissions(TE_t) = \alpha_0 + s(P_t^{NG}) + s(Q_t^E) + \mathbb{1}\{P_t^{NG} > med(P_t^{NG})\} * s(Q_t^E) \\ + s(HDD_t) + s(CDD_t) + s(Renewables_t) + s(Date_t) + \gamma D_m + \epsilon_t \end{aligned} \tag{2.2}$$

I control for electricity demand for two reasons. Most importantly, including demand will eliminate a major source of possible endogeneity from, e.g., macroeconomic conditions or weather conditions. If the recent economic downturn were correlated with the fall in natural gas prices, the results could be biased. The economic downturn would cause lower emissions through lower gas prices, but also through a decrease in electricity generation. Thus, the analysis could overstate the decrease in emissions caused by the decrease in the natural gas price. Similarly, if warmer summers increased gas prices and electricity consumption, they would cause bias. In this case, higher gas prices would be correlated with higher consumption and increased coal consumption. This could also cause the analysis to overstate the decrease in emissions caused by the decrease in the natural gas price.

The second reason that I include electricity demand is that over the medium-to-long run lower gas prices may cause lower electricity prices, increasing the quantity demanded of electricity. To the extent that medium run effects exist, including demand as a control allows me to isolate the first-order effect of gas prices on carbon emissions. While including demand directly in the model is unconventional, it is likely appropriate in the electricity sector. I assume that, in the short-run, demand is determined exogenously outside the model. This is reasonable because prices are generally not available in real time, thus fixing the quantity of electricity demand over short periods of time.

It is important to control for the level of renewable generation because renewable generation directly replaces conventional generation. Wind and solar patterns are seasonal and cause renewable electricity generation to also be seasonal. For instance, in the Western interconnection, winds are strongest around April and are generally much weaker in October. Therefore, renewable production peaks in April and is much lower in the late fall. Gas prices are also seasonal. Failing to control for the variation inherent in renewable electricity production would cause bias.

[33] For the spline on HDD and CDD I use three knots. The ambient temperature ranges are too small to allow for six knots.

While electricity demand and carbon emissions data are hourly and gas prices are daily, renewable electricity generation is only available at the monthly level. Renewable generation has very low marginal costs. As a result, it should always come before gas and coal in the dispatch order. On a day-to-day level, it is not very correlated with gas prices. Thus, the lack of granularity in the data will only cause very limited bias in my results.[34,35,36]

It is possible that failing to control for generator efficiency changes caused by changes in ambient temperature causes my results to be overstated. This could be the case, e.g., because hot days increase both carbon emissions and electricity demand or the gas price. Hotter days may cause generators to operate less efficiently, directly increasing gas usage and carbon emissions. They could also increase electricity demand (air conditioning) or the gas price (more demand for electricity). Controlling for ambient temperature will prevent this type of bias.

HDD and CDD are better suited to analyze temperature's effect on generator efficiency than raw temperature. A raw temperature average could disguise important temperature heterogeneity. For example, if the Western interconnection had a temperature of 68 F in all areas, this would result in a raw average temperature of 65 F, HDD of 0, CDD of 0, and very little effect on generator efficiency. However, it could also be the case that California is very hot and the rest of the west is very cold. Here, the population-weighted average temperature would still be 65 F, but the HDD could be 10 and the CDD could be 10. Generator efficiency would differ from the former case. Using HDD and CDD allows me to more accurately capture the effect of ambient temperature on carbon emissions through generator efficiency.

I include month-of-year dummies to control for residual seasonal variation not captured by my renewables data. This might arise because of seasonal generator maintenance. The flexible time trend is used to control for trends through time. In particular, the generation mix and international

[34] In section 2.7.1 I check the robustness of this assumption using hourly wind data that are available only in the Texas interconnection. In all years of my sample, wind, wood, and hydroelectric power are by far the largest sources of renewable power. For more, please see the EIA's Electric Power Monthly, table 1.1.A: `http://www.eia.gov/electricity/monthly/`.

[35] It is possible bias could result from cloudy periods. They are cooler, resulting lower electricity demand and lower solar-powered generation. The extent to which this causes bias is likely minimal. In 2013 (the year in my sample with the largest amount of solar generation), solar produced 9 GWh of power in the United States. Total US generation in 2013 was 4.1 million GWh, making solar responsible for less than 0.01% of total 2013 generation. For more, please see the EIA's Electric Power Monthly, tables 1.1 and 1.1.A: `http://www.eia.gov/electricity/monthly/`.

[36] It is also possible that hydroelectric generation is correlated with gas prices. This could happen because operators are able to produce electricity when gas prices are relatively high within the same month. This is unlikely because it requires that operators know whether near future prices will be higher or lower than current prices. Predicting future gas prices is very difficult. To the extent that this hydroelectric generation is correlated with gas prices, this would work against finding results in this paper. Because operators would be providing more renewable power when gas prices are high, the gas price spline would be flatter.

demand for coal are changing slowly over time. Including the time trend allows me to control for these changes. The time trend is likely sufficient to control for new capacity additions. While they are important, it is not the case that they are large relative to the existing generation stock; they increase the generation stock by roughly 6%.[37] Note that while I can control for the effect the changing generation mix has on the gas price spline, it does not preclude me from estimating its effects on carbon emissions. As discussed below, the effects of the generation mix are estimated by looking at the demand spline.

Marginal emissions, which new gas-fired plants are displacing, could vary with the gas price because the dispatch order of power plants adjusts as gas prices change. The interaction term, $\mathbb{1}\{P_t^{NG} > med(P_t^{NG})\} * s(Q_t^E)$, allows me to examine how high gas prices interact with the demand spline. The term is a "high gas price demand spline." It is constructed by finding the median gas price over my time period and creating a dummy if the gas price is above the median. I then multiply this dummy by a demand spline to allow for marginal emissions from demand to vary when the gas price moves above the median.[38]

Previous literature sometimes uses the ratio of the natural gas to the coal price as an independent variable. I prefer to omit the coal price because the coal price is in part determined by the natural gas price.[39] If natural gas were more expensive, demand for coal would be substantially higher. I aim to capture the total effect of increased natural gas supplies on carbon emissions through gas and coal prices - not only through the price of natural gas itself. I consider specifications with a coal price in Appendix B.5; results are similar to those using only gas prices.

The Durbin-Watson statistic suggests that autocorrelation may be an issue. I use Newey-West standard errors with seven lags. I choose seven lags for two reasons. First, it is a full week. It is possible that a firm's decision today (e.g., Tuesday) is correlated with Monday's decision, as well as the decision that they took on the previous Tuesday. Second, Greene (2012) recommends using the fourth root of the number of observations, which in this case is just under seven.

Aggregate calculations which combine results from several regressions use bootstrapped standard errors. I use block bootstrapping with 1,000 replications to mimic the possible autocorrelation

[37]If I use year-of-sample dummies instead of a time trend, the resulting price spline is qualitatively similar, but somewhat flatter. Gas prices are consistently high in the first two years, and consistently low in the last four years. The inclusion of year-of-sample dummies causes the estimation to struggle with the transition between the (pre-fracking) high gas price regime and the (post-fracking) low gas price regime.

[38]The median gas price in my sample is about $4/MMBtu. I also run the analysis using $6/MMBtu as the break point. Results are similar.

[39]Note that while fracking likely causes the coal price to exogenously change, the coal price is also driven by trends like changing international demand. My specification controls for trends over time, but directly including the coal price would not allow me to control for these trends.

in the data. Where possible, I have compared analytic standard errors with bootstrapped standard errors for accuracy. They are similar.

2.5.2 Decrease in Emissions from New Natural Gas Capacity

The key relationship when estimating the decrease in emissions from new natural gas capacity is between emissions and the electricity demand spline. The CEMS database allows me to directly calculate how much electricity was generated by newly constructed power plants, as well as the carbon that was emitted when generating the electricity. This power would otherwise have been produced by the old generation stock. Using the actual conditions at the time of power generation, I generate counterfactual emissions by increasing electricity demand and moving up along the demand spline by the amount of power that new plants are generating. A simple comparison between actual emissions from the newly constructed plants and marginal emissions from the counterfactual reveals the decrease in carbon emissions caused by these new facilities.[40,41]

I have aggregated generation from gas-fired plants constructed between 2010 and 2013. Figure 2.6 demonstrates that they have played an increasingly large role in US power generation. New plants are most important in the Eastern interconnection; their contributions in the Texas and WECC are considerably smaller. By the end of 2013, well over 10 GW of generation is supplied at any one time by these plants. On average, the US is generating about 450 GW of electricity – meaning that new gas-fired power plants made up about 3% of total US generation in 2013.

At this point the reader may be concerned that I do not control for new capacity additions, which could alter the shape of the gas price or demand splines. This is not likely to be a problem because the capacity additions are small relative to the existing capacity stock.[42] The time trend controls for much of these changes. Any residual bias would work against finding results in this paper because new gas-fired capacity allows for lower emissions levels and will flatten the gas price and demand splines.[43] A flatter gas price spline would cause emissions savings from lower gas prices to be (marginally) underestimated and my estimate of emissions savings from new plants

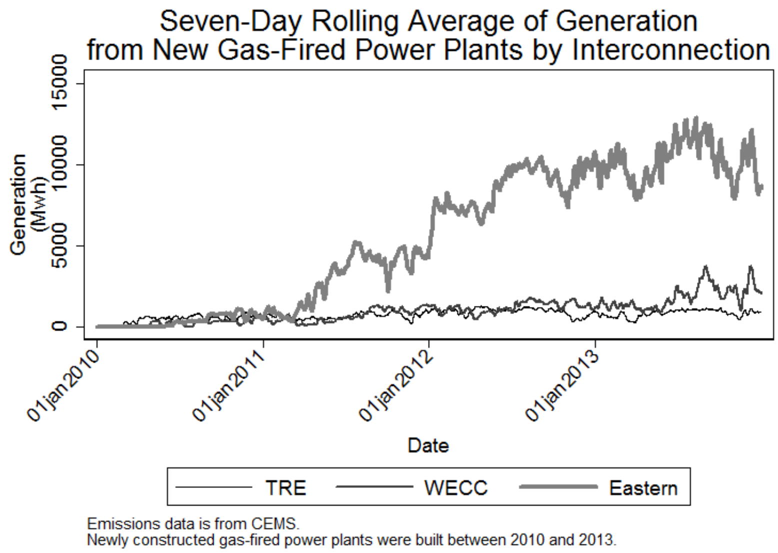
[40] This approach is similar to the one taken by Davis and Hausman (2014).

[41] My analysis does not consider the effect of delayed gas retirements or accelerated coal retirements due to low gas prices. These changes also yielded benefits to the extent that they shifted generation away from coal-fired sources. One important difference between adjusted retirement dates and new construction is that new construction will have a lifespan of several decades, while retirement adjustments only affect the marginal years surrounding retirement.

[42] Between 2007 and 2013, about 50 GW of natural gas capacity was added, relative to the existing generation stock of about 1,000 GW.

[43] New capacity is more efficient than older capacity and is profitable to run at higher gas prices than older plants. When gas prices drop to the point where switching between coal and gas starts to make sense, the new plants will be the first to be called upon.

Figure 2.6: Seven-Day Rolling Average of Generation from New Gas-Fired Power Plants



to also be lower.

For emissions reductions estimates, I only look at new capacity added between 2010 and 2013 (25.9 GW). Of this, I estimate roughly 65-85% was induced because of fracking (see Section 2.3).

In Section 2.7 I check the robustness of my primary specification to including daily wind generation or using additional Newey-West lags.

2.6 Results

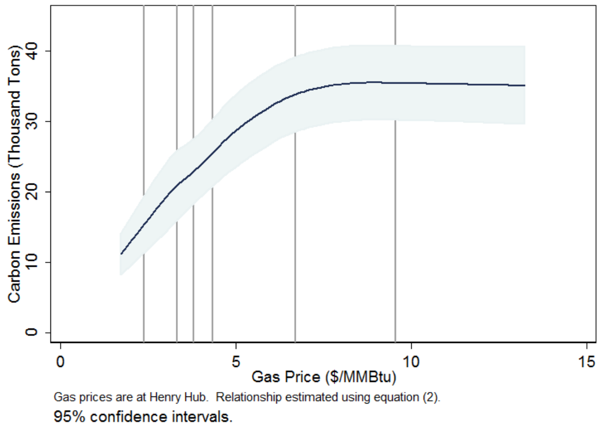
2.6.1 Decrease in Emissions from Switching

Figure 2.7 shows the relationship between the gas prices and carbon emissions in the Eastern, Western, and Texas interconnections as estimated using equation 2.2. The vertical lines represent the knot points in the splines.[44]

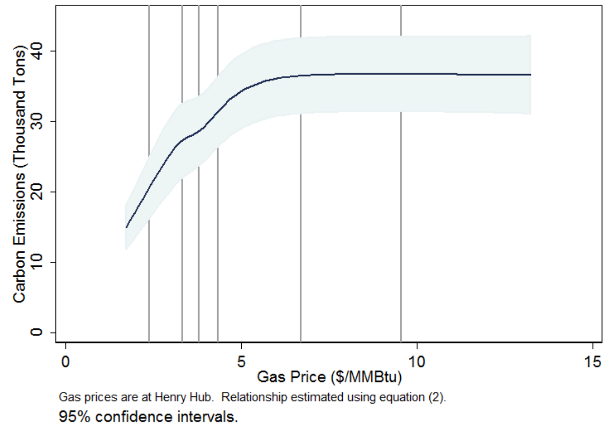
In all three interconnections, the gas price spline is strongly significant. The increase in carbon emissions from raising the natural gas price by $1 is highest when natural gas prices are low because coal and natural gas have similar marginal costs when gas is relatively inexpensive. Higher price sections of the splines have weaker (or non-existent) effects, as coal is cheaper than $7/MMBtu gas and it is also cheaper than $12/MMBtu gas. Note that the figures only include

[44] My primary specification does a very good job of predicting emissions, see Appendix B.3 for details.

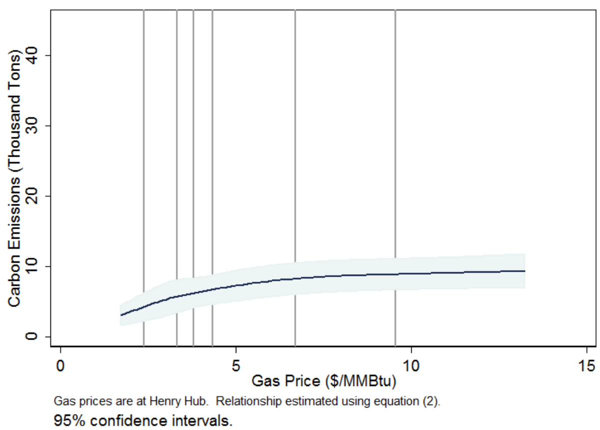
Figure 2.7: Hourly Relationship between Gas Prices and Carbon Emissions



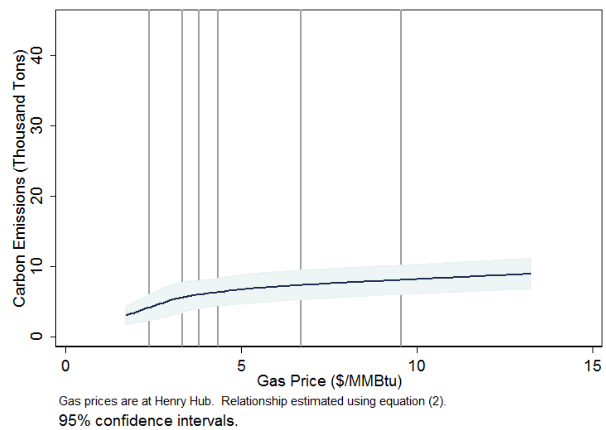
(a) Eastern: 2:00 AM (Off-Peak)



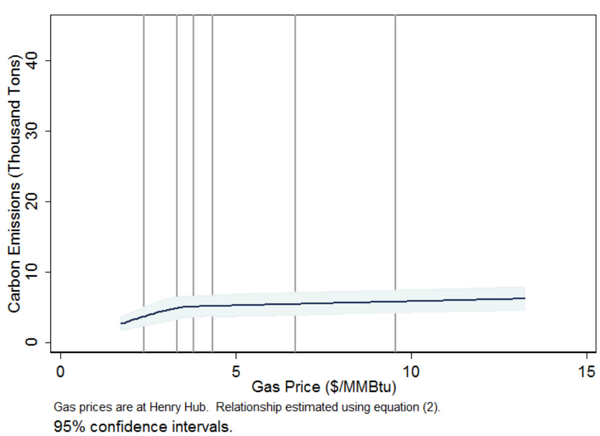
(b) Eastern: 6:00 PM (Peak)



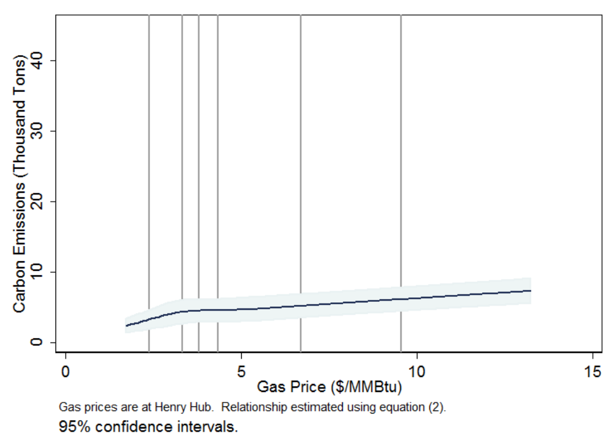
(c) Western: 2:00 AM (Off-Peak)



(d) Western: 6:00 PM (Peak)



(e) Texas: 2:00 AM (Off-Peak)



(f) Texas: 6:00 PM (Peak)

the effect of the gas price – fixed effects and other covariates (e.g., electricity demand) have been stripped out.

The Eastern interconnection has much higher levels of emissions because it is much larger than the other two interconnections. Additionally, the slope is much steeper at low gas prices because there is a lot of coal-fired generation that can be replaced in the Eastern interconnection, resulting in massive emissions savings. In contrast, due in part to their smaller sizes, the Western and Texas interconnections have less coal-fired generation.

In Figure 2.8, I plot seven-day average (rolling) emissions reductions if gas prices were at 2008 levels. Counterfactual emissions are constructed by taking each day's control variables as given, except the gas price is replaced by the gas price from the corresponding date in 2008. For example, the real price of natural gas on October 9, 2012 was \$3.00/MMbtu. On October 9, 2008, the real price of natural gas was \$6.72/MMBtu. In constructing counterfactual emission levels, I keep renewable production, electricity demand, and fixed effects at the levels on October 9, 2012, but use \$6.72/MMbtu as the gas price.[45] I choose 2008 because the gas prices from this year represent the natural gas market prior to the effect of fracking. Results using 2007 gas prices are qualitatively and quantitatively similar.

Counterfactual emissions are higher than actual emissions. This is exactly as expected – at higher gas prices, more coal is being burned. Emissions decreases are greatest in 2012, the year with the lowest gas prices. Decreases in 2010 are smallest (though still substantial), as this is the (post-2008) year when gas prices were highest (see Figure 2.3).

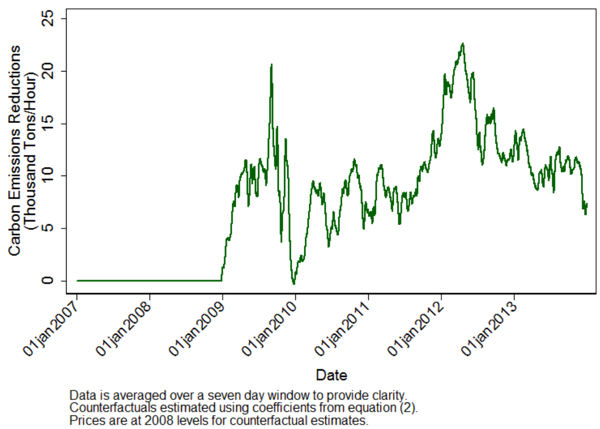
Table 2.4 details annual emissions decreases. All years show decreases that are substantial in magnitude. Depending on price fluctuations within a specific year, annual emissions reductions caused by low gas prices range between 9.7 and 22.0 thousand tons/hour (row [d]). This is between 3.4% and 7.7% of the 2008 total (s.e. of 0.4%).[46] On average, emissions have been 8.2% lower than 2008, and gas prices were directly responsible for a decrease of 5.0%. Lower gas prices are responsible for 61% of the total decrease. These decreases are larger in magnitude than has been previously estimated.

Remember that this estimate does not include possible increases in emissions due to lower electricity prices (and, consequently, higher electricity quantity demanded). I address what this demand response might look like in the discussion section. Additionally, the reductions discussed

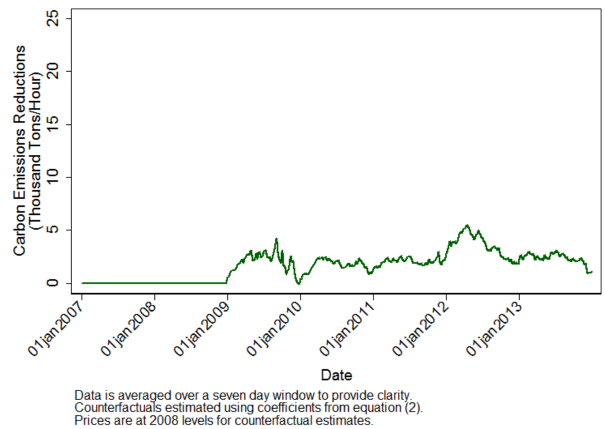
[45] I also adjust the interaction term.

[46] This estimate is a combination of effects in the three interconnections. As such, the standard error is calculated using block bootstrapping.

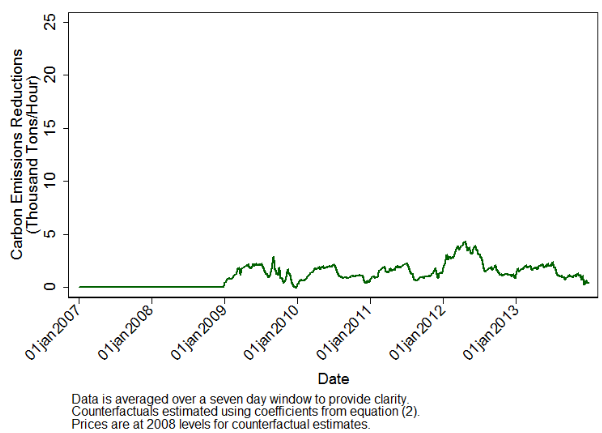
Figure 2.8: Emissions Reductions due to Lower Gas Prices



(a) Eastern Interconnection



(b) Western Interconnection



(c) Texas Interconnection

Table 2.4: Hourly Reductions in Carbon Dioxide Emissions

Hourly Reductions in Carbon Dioxide Emissions
By Year
(Thousand Tons of Carbon Dioxide/Hour)

		2008	2009	2010	2011	2012	2013
		[1]	[2]	[3]	[4]	[5]	[6]
Actual Emissions	[a]	284.7	261.3	278.3	266.0	249.0	253.0
Predicted Emissions using Primary Specification	[b]	284.3	261.9	277.6	266.5	249.2	252.7
Counterfactual Emissions (Using 2008 Gas Prices)	[c]		273.9	287.2	279.1	271.2	267.5
Emissions Reductions From Gas/Coal Switching	[d]		12.0	9.7	12.6	22.0	14.7
			(0.8)	(0.8)	(0.9)	(1.1)	(0.9)
Counterfactual Emissions (Without Newly Constructed Plants)	[e]			277.6	267.1	251.0	254.9
Reduction Caused by New Power Plants	[f]			0.1	0.6	1.8	2.1
				(0.0)	(0.0)	(0.1)	(0.1)
Counterfactual Emissions (With 2008 Gas Prices and Without New Plants)	[g]		273.9	287.3	279.7	272.9	269.4
Total Reductions from Low Gas Prices	[h]		12.0	9.7	13.2	23.7	16.7
			(0.8)	(0.8)	(0.9)	(1.1)	(0.9)

Notes:
[d] = [c] - [b]
[f] = [e] - [b]
[h] = [g] - [b]
Standard errors are estimated using block bootstrapping with 1000 replications.

in this subsection are restricted to reductions from switching between gas and coal-fired power plants. I now address the effect that new capacity had on emissions.

2.6.2 Decrease in Emissions from New Natural Gas Capacity

It is important to control for the exact level of demand (by running an hourly specification) when new gas-fired capacity is operating. Figure 2.9 illustrates this by showing the relationship between electricity demand and carbon emissions in all three interconnections at 2:00 AM (off-peak) and 6:00 PM (peak). Marginal emissions can vary based on the level of demand or the hour of generation.[47]

The difference in the demand splines is subtle in the Eastern interconnection (panels (a) and (b)). The 2:00 AM spline is a little steeper than the 6:00 PM spline. This is likely because coal-fired generation is more frequently used to meet marginal demand during off-peak hours than during peak hours.

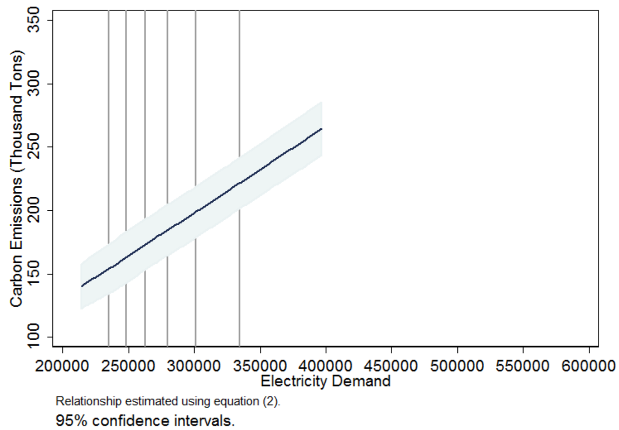
The importance of using a demand spline is most visible in the Western interconnection (panels (c) and (d)). When demand is around 50,000 MW, incremental demand causes very low incremental emissions (panel (c)). However, marginal emissions are higher at demand levels above 60,000 MW. This is largely because, in the Western interconnection, the marginal fuel switches from renewable hydro-electric generation to gas-fired generation as demand increases from very low levels to more moderate levels.

The Texas interconnection (panels (e) and (f)) shows the least variation in marginal emissions across hours and demand-levels, though the demand splines have some non-linearities. The majority of marginal generation is met by gas-fired plants, and changing marginal emissions are likely due to differences in generator efficiency. In particular, there is a slight increase in the slope of the demand spline when demand increases from low to moderate levels – levels when less-efficient gas-fired plants are running.

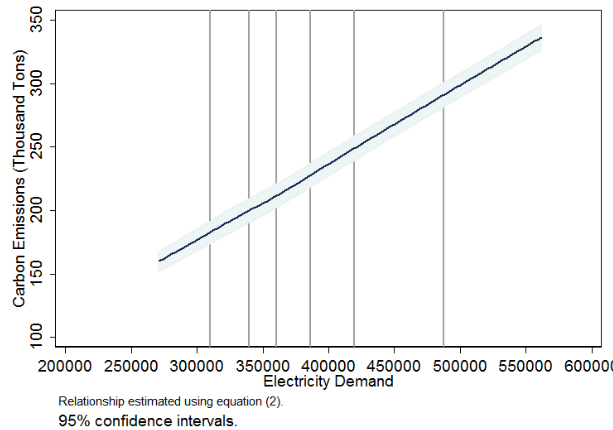
Figure 2.10 graphs seven-day rolling emissions reductions due to new gas-fired plant construction. Due to the lead time required to build a new gas-fired power plant, I only consider emissions reductions from plants that came online in 2010 or later. Strikingly, reductions are concentrated in the Eastern interconnection. This is largely true because the Eastern interconnection is the largest and has the most new plants. In 2012, the Eastern interconnection averaged 8.5 GW of generation at any time from new plants, while the other two interconnections each generated about 1 GW. Additionally, marginal emissions from incremental demand are higher in the Eastern interconnection

[47] Figure 2.9 does not incorporate the “high gas price demand spline.”

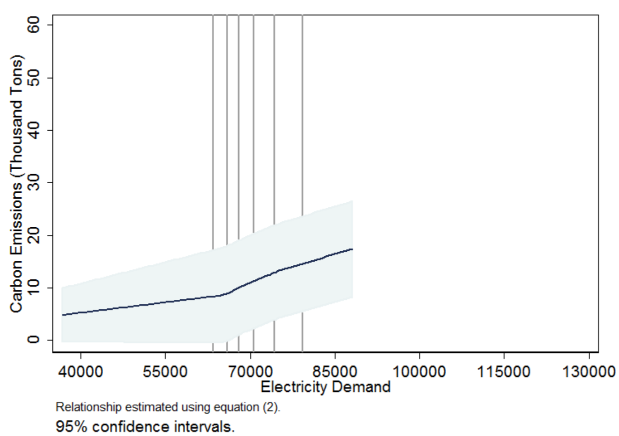
Figure 2.9: Hourly Relationship between Electricity Demand and Carbon Emissions



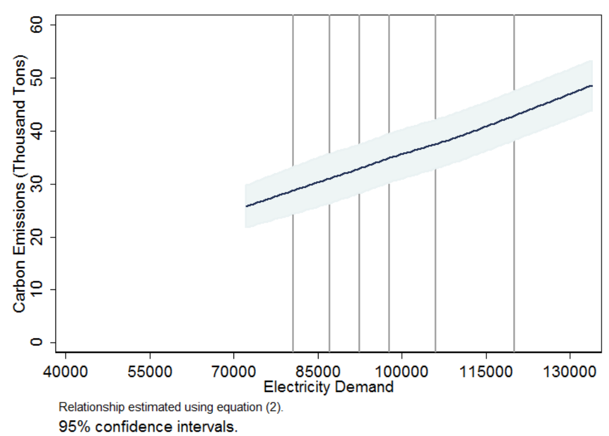
(a) Eastern: 2:00 AM (Off-Peak)



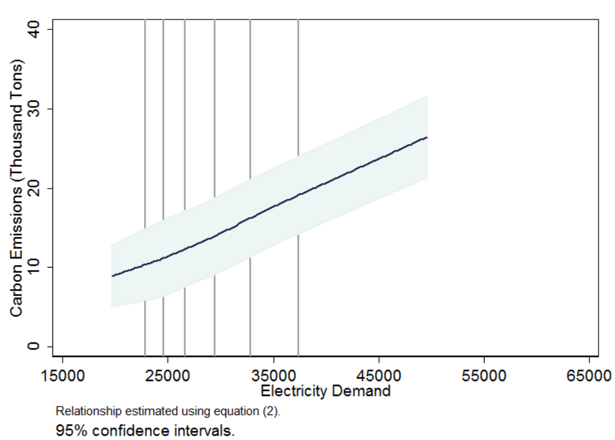
(b) Eastern: 6:00 PM (Peak)



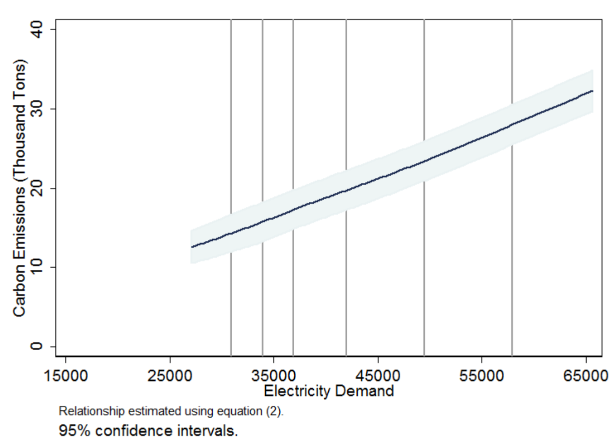
(c) Western: 2:00 AM (Off-Peak)



(d) Western: 6:00 PM (Peak)

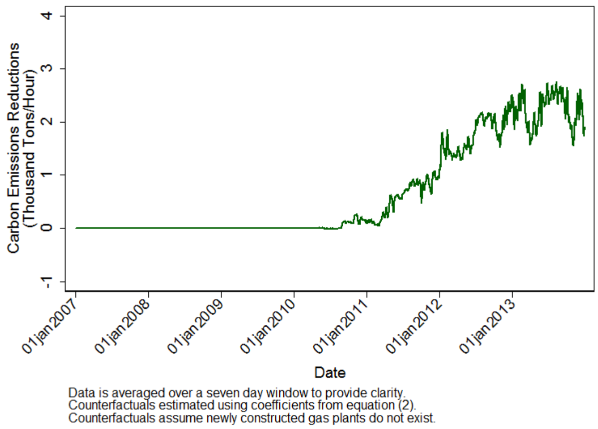


(e) Texas: 2:00 AM (Off-Peak)

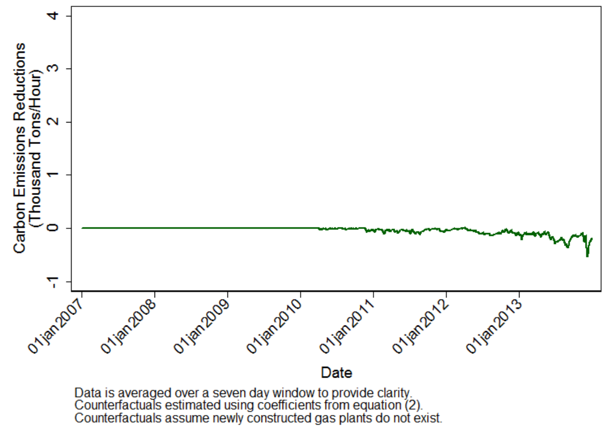


(f) Texas: 6:00 PM (Peak)

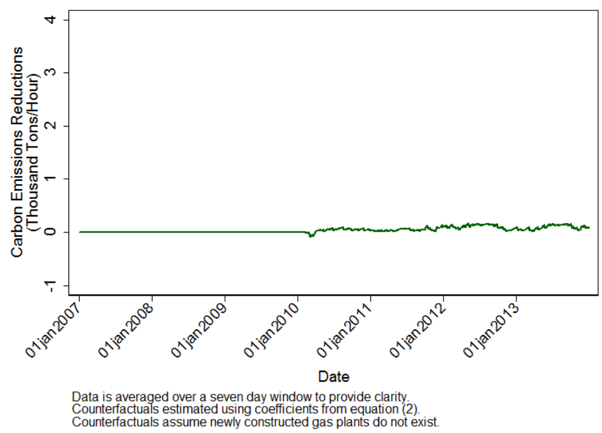
Figure 2.10: Seven-Day Rolling Emissions Reductions due to New Construction



(a) Eastern Interconnection



(b) Western Interconnection



(c) Texas Interconnection

because it is more reliant on coal-fired generation. New gas-fired generation in the Eastern interconnection is more likely to offset coal-fired generation, and corresponding emissions reductions will be larger in the Eastern interconnection.

Note that this was not a foregone conclusion. It could have been the case that the new gas-fired power plants were not running very frequently or were replacing similar gas-fired power plants. This could have led to no net emissions reductions. Additionally, Figure 2.10 shows several places where counterfactual emissions are actually lower than actual emissions.[48]

[48]In particular, the WECC shows some emissions *increases* at the end of 2013. In contrast to the other interconnections, most of the new plants in the WECC are single-cycle "peakers" that have low capital costs, but are not very efficient. When combined with the fact that marginal emissions are lowest in the WECC, emissions increases are possible. The magnitude of these increases is very small.

Table 2.4 details annual emissions decreases from newly constructed capacity (see row [f]). As expected, continued additions over time cause emissions reductions to grow over time. By 2013, hourly emissions savings averaged 2.1 thousand tons. This is 0.75% of the 2008 total. As detailed in section 2.3, I estimate about 65-85% of the 2.1 thousand tons/hour is directly attributable to lower gas prices. Because the capital stock is brand new, these gains will likely persist for years.

The reduction in emissions due to construction of new plants is less dependent upon low gas prices. Many new gas-fired power plants are very efficient. They fall below some coal plants in the dispatch order even when gas prices are moderate. Their combined-cycle technology can achieve efficiency of around 50%. In contrast, older plants are more likely to use single-cycle technology that only allows for efficiency around 33%.

2012 was warmer than average, which led to gas prices that are lower than average – and also lower than in 2013. Despite this, emissions reductions from newly constructed power plants continued to grow; 2013 emissions reductions were actually larger than those in 2012. A return to pre-fracking gas prices would be unlikely to negate all of the emissions gains from the new capital stock that has been built.

2.6.3 Combined Decrease in Emissions from Low Gas Prices and New Natural Gas Capacity

I now consider the combined effect of lower gas prices and new natural gas capacity. Specifically, I use my primary specification and adjust both the gas price to 2008 levels and electricity demand up as in the previous two sections (and make corresponding changes to the interaction term).

There are primarily two countervailing effects that determine how the interaction of low gas prices with new construction will affect carbon emissions. Lower gas prices, when combined with new plant construction, mean that it is likely new plants will run more frequently than they would if gas prices remained at higher levels. This would suggest that the combined effect should be larger. Working against this is that both changes, individually, might end up causing the same adjustments to the dispatch curve. That is, e.g., building a new gas-fired plant would cause it to displace a certain coal plant. If, instead, gas prices were lower, an existing plant might also displace the same coal plant. However, it is clear that one coal plant can only be displaced once. This would suggest that the combined effect should be smaller. However, it is possible that more than one coal plant is able to be displaced, allowing for positive synergies.

If I combine these changes and estimate a counterfactual where gas prices are at 2008 levels and no new gas-fired capacity was constructed, 2013 hourly emissions become 269.4 thousand tons of carbon/hour. This is close to 2008's levels of 284.7 thousand tons of carbon/hour. Gas prices and new gas fired capacity are responsible for a reduction of 16.7 thousand tons of carbon/hour (s.e. of 0.9). This is a substantial decrease of 5.9% from 2008 levels. Table 2.4 again details these changes by year (row [h]).

The combined effect of these two changes is slightly less than the sum of its parts. In 2013, lower gas prices reduced emissions by 14.7 thousand tons of carbon/hour, while newly constructed capacity reduced emissions by 2.1 thousand tons of carbon/hour. However, total reductions of 16.7 thousand tons of carbon/hour are 0.2 thousand tons/hour less than the 16.9 thousand tons/hour that are the sum of the parts. This suggests that potential synergies are outweighed by the inability to displace the same dirty plant twice. Note that this effect is relatively small.

2.7 Robustness Checks

I consider several alternative specifications to alleviate concerns about my results being due to misspecification or chance. In this section I focus on including daily wind generation and using additional Newey-West lags. In Appendices B.4 and B.5 I consider using alternative gas prices or a gas/coal price ratio.

2.7.1 Inclusion of Daily Wind Generation Data

It is possible that failing to control for wind causes my results to be overstated. This could be the case because windy days decrease both carbon emissions and the gas price. As more wind power is generated, less fossil-fuel generation is needed. As a result, carbon emissions will drop, as will the prices of fossil fuels. Daily wind generation is only available in the Texas interconnection. Texas has the most wind capacity as a percentage of electricity demand (the Western interconnection has about two-thirds as much and the Eastern interconnection has about one-third as much). Thus, any effects of wind generation should be smaller in the other two interconnections.

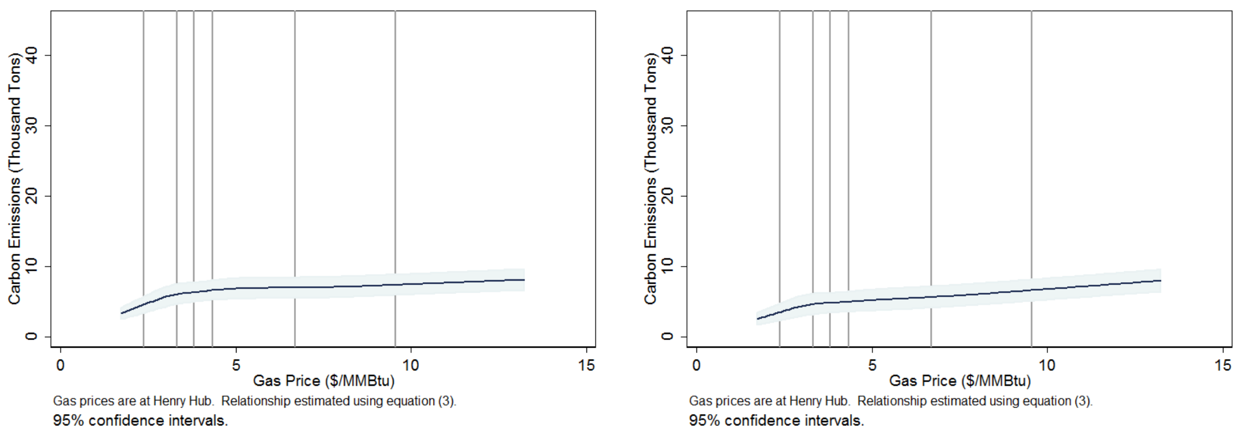
Specifically, I estimate the following regression[49]:

$$TE_t = \alpha_0 + s(P_t^{NG}) + s(Q_t^E) + \mathbb{1}\{P_t^{NG} > med(P_t^{NG})\} * s(Q_t^E) + s(Renewables_t) + s(Date_t) + \gamma D_m + \mathbf{s}(\mathbf{Wind_t}) + \epsilon_t \quad (2.3)$$

[49] I also include two dummies to account for data irregularities. Most of the data is from a colleague, while about 2% is directly from ERCOT. Additionally, approximately 2% of the data remains missing. I fill missing data with the observation from 24 hours prior to preserve the structure of the dataset.

The resulting splines and counterfactual emissions reductions are similar. Figure 2.11 shows two splines that are comparable to those in Figure 2.7. There are almost no differences between the two figures. This translates into limited change in the estimates. For example, my primary specification estimates that Texas carbon emissions in 2013 were 1.46 thousand tons/hour lower than they would have been with higher gas prices (s.e. of 0.21). By including wind in my specification, this estimate actually increases to 1.64 thousand tons/hour (s.e. of 0.20). I interpret these estimates as essentially the same. This suggests that the estimates presented in this paper do not have a large bias due to the failure to include daily wind generation. Any potential bias would be mitigated in the (larger) Eastern and Western interconnections by the fact that wind generation is a substantially smaller portion of the generation portfolio.

Figure 2.11: Texas Interconnection Gas Price Splines Including Wind Generation



(a) Texas: 2:00 AM (Off-Peak)

(b) Texas: 6:00 PM (Peak)

2.7.2 Additional Autocorrelation Lags

I now consider how the standard errors would change if I allowed for additional periods of autocorrelation. While my preferred specification allows for one week of autocorrelation, this section allows for one month of autocorrelation. For estimates that are composites of multiple regions or multiple hours, this means that I use larger blocks in my block bootstrap.

Table 2.4 displays estimates using seven days of autocorrelation. For example, my combined estimate for 2013 of 16.7 thousand tons of carbon/hour has a standard error of 0.9. Increasing the allowed autocorrelation to one month increases the standard error to 1.4. The estimate of reductions caused by new plants is 2.1 thousand tons of carbon/hour, with a standard error of 0.062. Using additional lags increases the standard error to 0.064. All estimates remain significant at the 1%

level.

2.8 Discussion

2.8.1 Estimated Value of Offset Emissions

I now estimate the economic value of reduced emissions. This is important because it allows one to better understand the magnitude of the benefits. The US Government recently provided an updated estimate of the social cost of carbon; it is currently about $35/ton (Interagency Working Group on Social Cost of Carbon, 2013). I estimate that lower gas prices offset about 14,700 tons/hour of carbon emissions in 2013. At current valuations, this is worth about (365 * 24 * 14,700 * 35 =) $4.5 billion. In 2013, newly constructed gas-fired power plants reduced carbon emissions by about 2,100 tons/hour. This is worth about (365 * 24 * 2,100 * 35 =) $0.65 billion. My estimates from Section 2.3 suggest that between $0.43 and $0.56 billion of the $0.65 billion (65-85%) is due to fracking. Combined, I estimate the 2013 decrease in carbon emissions is worth about $5.1 billion. Most of this benefit is a pure externality, as the market only prices carbon in the RGGI states and, starting in 2013, California.[50]

The value of the emissions reductions varies between 2009 and 2013. The least valuable year was 2010, with offsets worth about (365 * 24 * 9,700 * 35 =) $3.0 billion. The most valuable year was 2012, with offsets worth about (365 * 24 * 23,700 * 35 =) $7.3 billion.

2.8.2 Demand Response

It is important to recognize that additional electricity demand and emissions may have been induced by lower gas prices (which led to lower electricity prices). Indeed, Linn, Muehlenbachs and Wang (2014) find that natural gas and electricity prices have a positive and causal relationship. Estimating the demand response to lower electricity prices is difficult. In 2008, average electricity prices were 10.32 cents/kilowatt-hour, with 7.07 cents/kilowatt-hour coming from generation. The 2008 AEO projected that the average electricity price in 2013 would be 0.45 cents/kilowatt-hour lower, with a 0.61 cents/kilowatt-hour decrease in generation costs (transmission and distribution costs were projected to rise). This projection was made prior to information about the decline in gas prices. In reality, electricity prices fell by 0.53 cents to 9.79 cents/kilowatt-hour in 2013. Generation costs fell by 1.28 cents/kilowatt-hour. That is, electricity prices fell by (0.53 - 0.45 =

[50]For comparison, I use EIA data to estimate that generation, transmission, and distribution costs decreased by between $3 and $27 billion because of lower gas prices. My estimation methodology is analogous to the way I estimate potential demand response in the following section.

) 0.08 cents/kilowatt-hour more than expected, and generation costs fell by (1.28 - 0.61 =) 0.66 cents/kilowatt-hour more than expected.

Determining whether the 0.08 cents/kilowatt-hour or the 0.66 cents/kilowatt-hour number is appropriate to ascribe to lower gas prices is also challenging.[51] The difference is due to increased transmission and distribution costs, some of which could have resulted from changing generation patterns. I will use these as the bounds of the potential demand response. A decline of 0.08 cents/kilowatt-hour is equivalent to a decrease of 0.78%, while a decline of 0.66 cents/kilowatt-hour is equivalent to 6.44%.

Using an elasticity of -0.3, these estimates suggest that electricity demand has increased by between 0.23% and 1.93%. I can now use my primary specification (2.2) to estimate the rebound effect. For the lower bound in 2013, I estimate lower electricity prices prompt increases in demand that cause carbon emissions to increase by 0.6 thousand tons/hour (s.e. of 0.005). For the upper bound, I estimate carbon emissions increase by 5.2 thousand tons/hour (s.e. of 0.04). This is between 4% and 36% of the estimated 14.7 thousand tons/hour of emissions decreases caused by lower gas prices.[52]

2.8.3 Other Environmental Considerations

Even considering potential demand response, it is clear that new gas supplies have decreased carbon emissions in the electricity sector. However, it is also the case that drilling for natural gas can have deleterious effects on the environment. In particular, if drillers are careless, leaking methane can offset many of the carbon emissions gains. President Obama's recently announced methane regulations demonstrate that this is being taken seriously by the federal government. The extent to which methane leaks are eliminated will largely determine whether or not fracking has a net positive effect on American greenhouse gas emissions. For a meta-analysis and overview of the shale gas life-cycle greenhouse gas emissions literature, see Heath et al. (2014).

2.9 Conclusion

Instead of declining, as projected several years ago, US natural gas production has dramatically increased over the last five years. This happened in part because the federal government has allowed it to. Governments (excluding New York's) have largely refrained from imposing regulations that

[51]This analysis assumes that consumers respond to average prices as Ito (2014) demonstrates. I use 2012 prices for these estimates.

[52]To the extent that other factors changed, such as increased Chinese demand for coal, these bounds may prove to be insufficiently narrow.

would seriously curb hydraulic fracturing for natural gas. I estimate that, depending on the year, 2009-2013 electric sector carbon emissions have decreased by between 3.4% and 8.3% as a result of lower gas prices. Lower gas prices have likely been a moderate contributor to the decrease in American carbon emissions over the last five years.

In many industries, capital stock is an important determinant of capability. It evolves slowly over time, even when large shifts have occurred within a sector. Investments in natural gas capacity will likely continue, so long as natural gas prices remain low. Recent government regulations intended to cut carbon make coal power plants even less desirable, increasing the amount of natural gas-fired capacity to be added in the near future. These plants will be with us for many years. Similarly, investments that were made decades ago in telecommunications, mass transportation, sewer lines, and interstates remain in regular use today, often way past their life expectancy.

Chapter 3

Airline Competition, Oil Price Pass-Through, and Carbon Taxes (With Anirudh Jayanti and Andrew Usher)

3.1 Introduction

Cost pass-through – the rate at which a cost change translates to a price change – is an important concept in many areas of economics, ranging from industrial organization (e.g., the effect of a merger) to international trade (e.g., exchange rate pass-through) and public finance (e.g., tax incidence). In particular, energy cost pass-through has important implications for environmental policy. In this paper, we use the Australian airline industry to study how fuel cost shocks are passed through to airfares and how the pass-through rate depends on the competitiveness of the route. Airlines are an intriguing setting to study pass-through due to the presence of many local "markets" (i.e. routes) existing simultaneously. Furthermore, from a policy standpoint, the airline industry is a significant emitter of carbon dioxide and is therefore an important part of the debate on strategies to address global climate change.

Fuel costs make up a significant portion of airlines' total costs and could potentially be passed on to consumers in the form of higher airfares. Additionally, the level of competition on a route likely affects how much of the cost is passed on. Pass-through in this context is important because the effect of higher fuel costs should be equivalent to the effect of a carbon tax (Cullen and Mansur, 2014). The pass-through rate we find could then help policymakers evaluate the likely effects of a carbon tax on the airline industry. Finally, by evaluating how pass-through differs with the number of competitors on a route and the type of product, we can shed light on the nature of competition in this industry.

This paper uses a novel dataset consisting of Australian airfare data from Sabre Corporation and capacity data from the Official Airline Guide (OAG). The airfare data is monthly, an important advantage over the publicly available, quarterly Databank 1B US data. The capacity data tells us exactly which aircraft each carrier used on each of its routes; we use this to construct a measure of fuel costs that is specific to each carrier and route. Our main empirical finding is that the pass-through rate is very high – around 100% for monopoly routes – and increases with additional competitors.

This paper has two main contributions. First, we estimate fuel cost pass-through in an industry characterized by differentiated products and imperfect competition. Because each firm offers multiple products (e.g. non-stop vs. one-stop flights), there is potential for pass-through to be different across products within the same firm. In fact, we should expect this to be the case due to different demand elasticities for different products. Relative to homogeneous-product industries, firms' ability to discriminate in this way could, for example, improve welfare by passing more of the cost on to relatively inelastic consumers. Because only the inelastic consumers are faced with a higher price, quantity supplied will be distorted less than if firms were unable to discriminate. These issues have been understudied in the empirical pass-through literature.

Second, we evaluate the heterogeneity of pass-through across different routes. We study how the pass-through rate varies by how competitive the route is. Theoretically, the relationship could go either way. A simple model of linear demand and constant marginal cost tells us that pass-through is 100% in perfect competition and 50% for a monopolist. If we further assume symmetric Cournot competition for the oligopoly case, pass-through rises monotonically from the monopoly rate to the perfectly competitive rate as the number of firms grows. On the other hand, other demand forms such as CES can have pass-through that is greater than 100% and decreasing with competition. This is therefore an empirical question. We find that pass-through increases with competition in the airline industry.

In answering these questions, our paper contributes to several literatures. A large literature in industrial organization studies market structure in the airline industry (Borenstein (1989); Berry (1992); Goolsbee and Syverson (2008); Ciliberto and Tamer (2009)). These papers estimate the impact of market dominance, entry, and the threat of entry. Another strand of this literature considers welfare effects of shorter-run decisions such as pricing, capacity, and codeshare agreements (Berry and Jia (2010); Armantier and Richard (2008)). Finally, there is a large literature that studies how price discrimination varies with competition in the airline industry (Borenstein and Rose (1994); Gerardi and Shapiro (2009); Chandra and Lederman (2015)). None of these papers consider fuel cost pass-through.

There has recently been renewed theoretical and empirical interest in pass-through. Weyl and Fabinger (2013) theoretically characterize several principles of pass-through under different market structures. Their main finding in general cases of imperfect competition is that pass-through depends crucially on the curvature of demand. Our reduced-form analysis imposes little structure on the curvature of demand, so we are not pre-determining pass-through with restrictive functional form assumptions. Unfortunately, this also means we cannot formally say anything about welfare, since in the Weyl and Fabinger (2013) framework this depends on both pass-through and a conduct parameter.

On the empirical side, Agrawal, White and Williams (2017) use Weyl and Fabinger (2013)'s framework to estimate tax incidence and competition in the US airline industry. Fabra and Reguant (2014), Miller, Osborne and Sheu (2017), and Ganapati, Shapiro and Walker (2016) study energy cost pass-through in homogeneous product industries (electricity, cement, and a subset of manufacturing, respectively). All three studies find high rates. Ganapati, Shapiro and Walker (2016) find a pass-through rate of 70% for manufacturing. Fabra and Reguant (2014) find that over 80% of emissions cost shocks are passed on to electricity prices. Miller, Osborne and Sheu (2017) find that fuel costs are more than fully passed on to cement prices.

Some studies have examined the relationship between pass-through and competition. Miller, Osborne and Sheu (2017) find that competition reduces pass-through in the cement industry. They explain this result using a symmetric oligopoly model with log-convex demand based on Weyl and Fabinger (2013). Ganapati, Shapiro and Walker (2016) compare several manufacturing industries and find that pass-through appears to be greatest in the least-competitive industry they study. On the other hand, Cabral, Geruso and Mahoney (2014) find that pass-through increases with competition in the Medicare Advantage insurance market. Given that different studies have found different results, the relationship between pass-through and competition is an empirical question. We use simulations and specific features of the airline industry to explain our finding that pass-through increases with competition.

Pass-through has also been studied extensively in contexts other than energy. In the international trade literature, Goldberg and Hellerstein (2013) estimate a structural model of the beer industry to explain incomplete exchange-rate pass through. This paper studies a differentiated product industry, but only in a single market (Chicago), whereas our setting allows us to look at heterogeneity across markets.

A separate literature at the intersection of industrial organization and environmental economics finds that the welfare effects of environmental regulation can be quite different in imperfectly competitive markets (Ryan (2012); Fowlie, Reguant and Ryan (2016)). These papers consider the cement industry; by contrast, we study the airline industry which is also a major source of emissions that are projected to grow rapidly.

The rest of the paper is organized as follows. In Section 3.2, we provide background on the Australian airline industry. In Section 3.3, we describe our data. Section 3.4 discusses the empirical model. Section 3.5 presents our main results, as well as a series of robustness checks. Section 3.6 concludes.

3.2 Background: Australian Airline Industry

This paper focuses on the domestic Australian air market. Figure 3.1 provides an overview of Australia's major airports. Australia is a particularly nice setting for our study because it is a relatively isolated country, decreasing the potential for outside events to substantially influence it.[1] Flying is the obvious way to travel between many of Australia's cities due to the lack of viable alternative transportation options. Many highways do not have frequent rest stops like those in the United States, and the rail network is much less developed than Europe's. Additionally, airline data are available at a higher frequency than similar publicly-available US data. Finally, Australia passed and repealed a carbon tax during our sample period. This policy change provides us with some additional variation and confirms that our results are relevant in the context of a carbon tax.

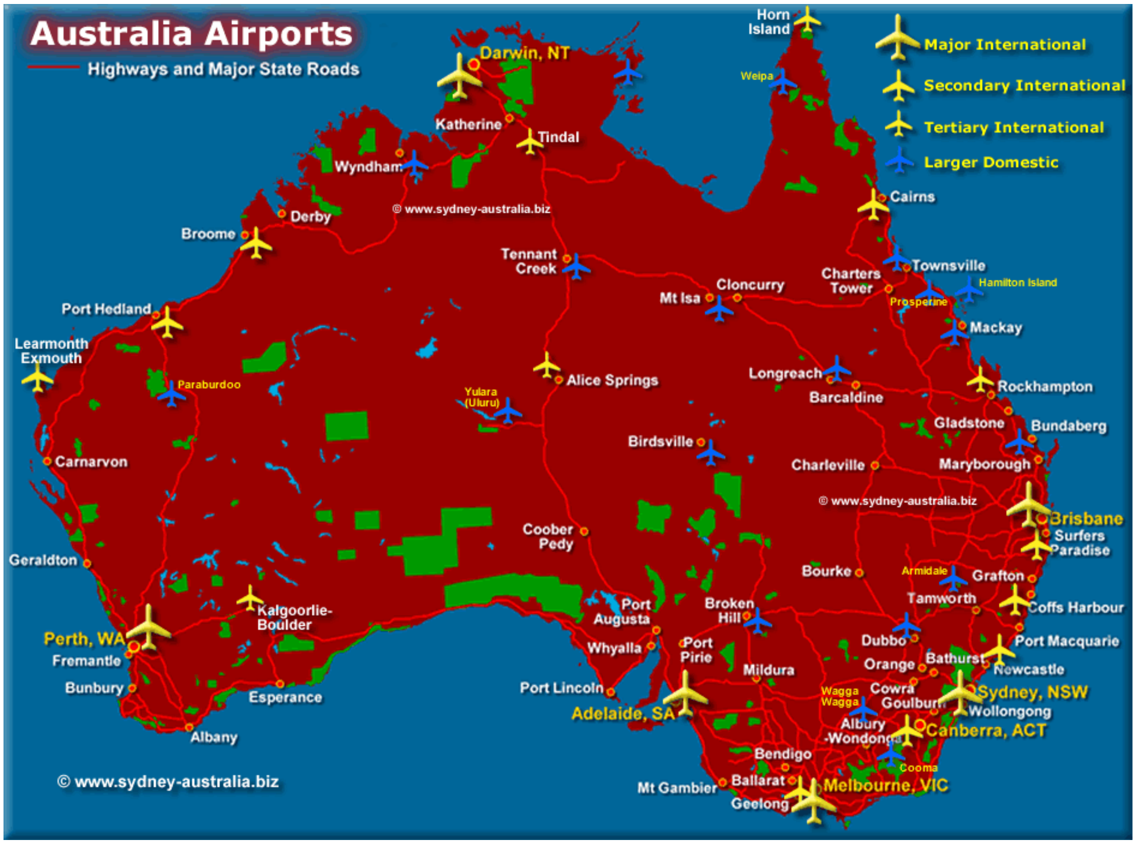
Australia's air market has two major full-service airlines – Qantas and Virgin – as well as a set of regional competitors and low-cost airlines. The regional and low-cost airlines have varying degrees of independence from the full-service airlines. While some operate fully independently, others have been purchased by Qantas or Virgin and become wholly-owned subsidiaries and/or are members of the same airline alliance. The latter relationship allows cross-marketing across each other's flights.[2] Table 3.1 shows each major airline's market share during our sample period.

Jet fuel is one of airlines' main costs and can constitute up to 30% of total costs. Jet fuel is a refined product that is made from crude oil. Consequently, its price is closely tied to the price of crude oil. The ease of transporting oil and its products means that prices are determined in world markets. World oil markets are very thick and no individual participant can substantially affect prices. We therefore treat jet fuel as an exogenous cost that airlines have no control over.

[1] For example, air travelers in the Netherlands have outside options at Dusseldorf Airport (Germany) and Brussels Airport (Belgium).

[2] For example, one could purchase from Qantas a flight that is operated by Northern Air Cargo.

Figure 3.1: Map of Australian Airports



Notes: Reprinted with permission from IA Connections. The original map is available at http://www.sydney-australia.biz/maps/australia/australiaairportsmap.php. Some small airports are not pictured, though they are present within our data.

Australia introduced carbon tax legislation in 2011. The legislation was subsequently passed and took effect on July 1, 2012. It was later repealed on July 17, 2014, but the repeal was backdated to July 1, 2014. We will treat July 1, 2012 through July 1, 2014 as months that airlines were expecting to pay the carbon tax. While most industries paid a tax of just under \$25AUD/ton, jet fuel was taxed at 6 Australian cents per liter. This tax was equivalent to roughly \$21.50/ton of carbon.[3]

Airlines have two primary ways to respond to changing jet fuel costs. Adjusting airfares is the easiest and most flexible way that airlines adapt – many airfares are changed multiple times per week. Airlines can also alter the set of routes that they offer, as well as the frequency with which each route is offered. However, this adjustment process generally takes considerably longer; it is

[3]This is within the range of estimates of the social cost of carbon, albeit at the lower end.

Table 3.1: Market Shares of Major Airlines

	Market Share	Total Passengers (Millions)
Qantas	38.57	152.72
Virgin	28.05	111.06
Jetstar	20.26	80.23
Tiger	5.13	20.31
Cobham Aviation	2.73	10.83
Regional Express	2.20	8.73
Skywest	0.65	2.58
Airnorth	0.48	1.90
Alliance	0.36	1.42
Skytrans	0.33	1.32

Notes: There are ten different airlines that operate with some degree of independence during our sample. Qantas and Virgin are the major full-service airlines. Jetstar is a low-cost wholly-owned subsidiary of Qantas. It competes with Qantas on many routes. Tiger was an independent low-cost airline until 2013 when Virgin purchased 60% of the airline. Virgin purchased the remaining 40% in 2014. The other six airlines are regional players that focus on less-serviced routes.

relatively expensive for airlines to adjust flights that they have already sold tickets for. We therefore focus on airlines' price responses to changing fuel costs.

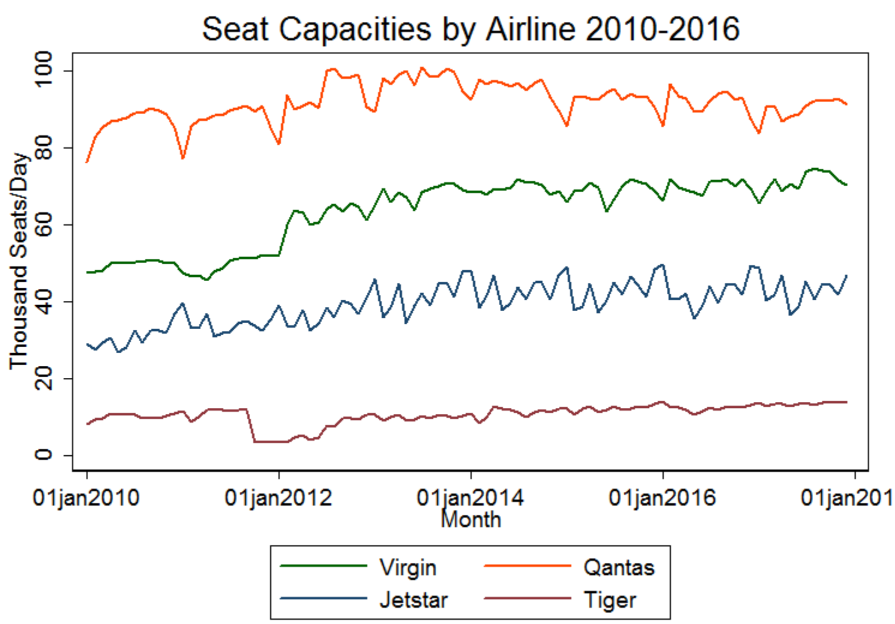
3.3 Data Description & Summary Statistics

Our airline price and passenger data was purchased from Sabre Corporation. It consists of detailed monthly level data for all domestic air travel in Australia from 2010 through 2015.[4] The raw data is aggregated so that everybody who flew on a given airline during a given month for an exact one-way itinerary in a cabin class is combined. For example, all passengers who flew discount economy on Qantas from Sydney to Perth via Canberra in January 2010 is one observation. For each observation, we see the number of passengers and total revenue in the month *traveled* rather than the month purchased (we explain below how we handle this constraint). The full sample consists of many routes with inconvenient layovers and very few passengers. Consequently, we drop observations where the airline has less than 1% market share for a route. These dropped observations are likely not viable competitors for most passengers. Total fare data is inclusive of all taxes, fees, and surcharges.

[4] Having the universe of tickets vs. a subsample is another advantage of our data over the DB1B.

Airline capacity and aircraft data was purchased from OAG. It contains the complete schedule of all domestic flights within Australia during our time period. We have data on the frequency, time, aircraft, and number of seats available for each route. Figure 3.2 shows capacity in terms of thousands of seats per day for the four largest carriers.

Figure 3.2: Seat Capacities by Airlines



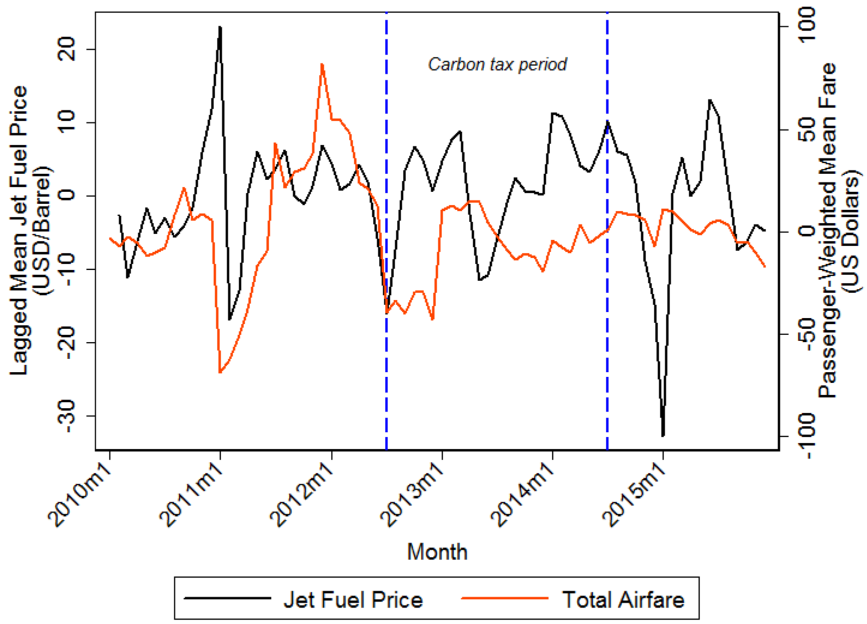
We construct the average cost of jet fuel (P_{irt}^{JF}) for each carrier-route-month observation by multiplying three terms together,

$$P_{irt}^{JF} = P_t^{Barrel} * Dist_r * FuelEff_{irt} \tag{3.1}$$

the average price of a barrel of jet fuel (P_t^{Barrel}) in month *t*, carrier *i*'s average fuel efficiency for a route ($FuelEff_{irt}$) in month *t* on route *r*, and distance ($Dist_r$) on route *r*. The average cost of jet fuel (P_{irt}^{JF}) varies with all three terms; different aircraft can have large differences in fuel efficiency, longer routes require more fuel than shorter routes, and the price of jet fuel varies over our time period. We now briefly discuss the source and construction of each term.

We collect daily jet fuel price data (P_t^{Barrel}) from Platts (accessed via Bloomberg) at the US Gulf Coast and average it by month. Because we do not know the date each ticket is sold, we assume that the average ticket is purchased one month in advance and lag our fuel prices by one month.[5] As discussed above, because oil is traded on global markets, we assume that the Gulf Coast price is a good proxy for the price Australian airlines face.[6] Jet fuel prices ranged between 45 and 163 US cents per liter during our sample period. Figure 3.3 plots average airfares and jet fuel prices. Both series are de-trended by year (to account for macro trends) and month (to account for seasonality). Total airfare tracks jet fuel closely for most of the sample.

Figure 3.3: Airfares and Jet Fuel Prices



Notes: Airfares are averages across Australia for a given month. Jet fuel prices are from the US Gulf Coast.

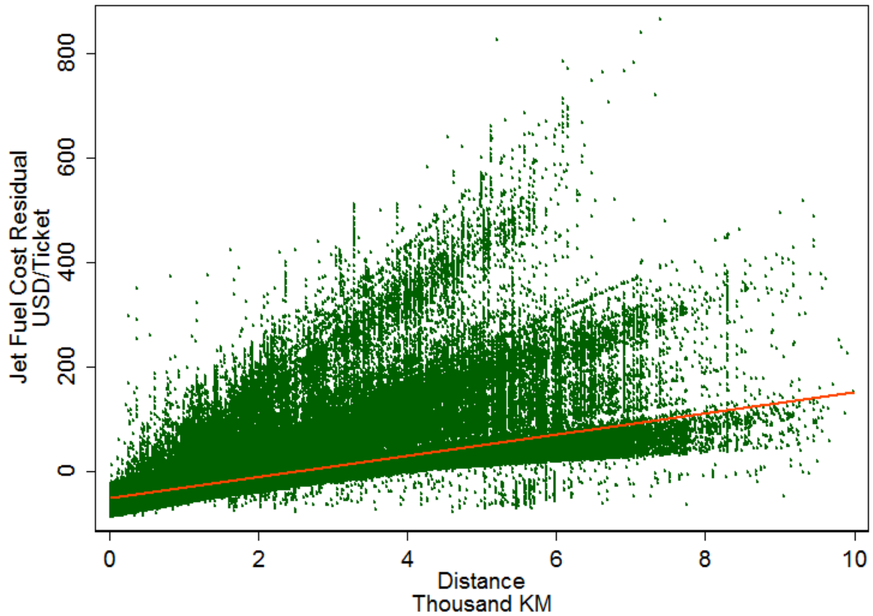
Distance ($Dist_r$) is one of the variables provided in our data from Sabre Corporation. Distance varies across and within routes – one-stop and non-stop itineraries will have different distances – and is an important source of fuel cost variation. Figure 3.4 provides a graphical representation of the variation from distance. The y-axis denotes the residual from a regression of jet fuel costs on month-of-sample fixed effects, while the x-axis shows each observation's distance. Distance is positively and meaningfully correlated with our fuel cost estimates.

Finally, for each observation we assign an average fuel efficiency ($FuelEff_{irt}$) based on the

[5] Our results are robust to varying the advance purchase assumption. See Table C.2 in Appendix C.1.

[6] Additionally, using Gulf Coast prices is advantageous because these prices are heavily correlated with Australian prices but not with Australia-specific demand shocks.

Figure 3.4: Jet Fuel Costs and Distances



Notes: This is a scatterplot of estimated jet fuel cost residuals and itinerary distances. Residuals are calculated after regressing jet fuel costs on month-of-sample fixed effects. The orange line is a line of best fit and demonstrates that the residuals are positively correlated with itinerary distance. Distance is positively and meaningfully correlated with our fuel cost estimates.

aircraft flying that route. Fuel efficiency data for each aircraft are gathered from internet research.[7] Fuel economy data varies widely by type of aircraft. Generally, larger aircraft use less fuel per seat-mile. Additionally, there can be large differences within aircraft class. For example, the 98-seat Embraer E-Jet-190 uses 3.81 L/100km/seat while the 82-seat Bombardier Dash 8 Q400 uses 2.79 L/100km/seat.

Table 3.2 summarizes our data at a slightly aggregated level: we combine passengers with different intermediate stops and passengers in different cabin classes. An observation is at the month-airline-route level (e.g., all seats on Qantas-operated flights from Sydney to Melbourne in January 2012, including routes with stops).

Recall from the discussion in Section 3.2 that the two dominant airlines – Qantas and Virgin – partly or wholly own regional airlines. In particular, Jetstar is a wholly-owned subsidiary of Qantas for the entire sample period. Consequently, we code any Jetstar flight as being operated by

[7] Wikipedia aggregates and cites estimates for many aircraft. Additional data is gathered for aircraft not listed there.

Table 3.2: Air Travel Summary Statistics

	Full Sample		Non-Stop Flights	
	mean	st.dev.	mean	st.dev.
Passengers	2760.32	10422.70	4466.20	11153.81
Average Total Fare (USD)	347.47	195.38	224.15	96.50
Average Fuel Cost (USD)	36.78	30.02	20.39	30.07
Distance ('000 KM)	1.79	1.43	0.87	1.28
Average Capacity Factor	0.69	0.12	0.70	0.16
Jet Fuel Price (US Cents/Liter)	113.46	26.22	113.34	26.07

Notes: Observations are at the month-route-airline level in the left two columns. For example, one observation is all passengers that flew from Sydney to Melbourne in January 2010 on Qantas. There are 112,965 total observations. The right two columns exclude data that has more than one stop.

Qantas; we do not want to label Jetstar as a competitor for Qantas because Qantas likely internalizes the profit impacts of Jetstar's pricing. Similarly, Tiger Airways was fully owned by Virgin after October 2012; we therefore code any Tiger Airways flight after that date as being operated by Virgin.

Code-sharing is common in Australia, as it is in the United States. For the purposes of estimating pass-through, we are interested in the identity of the operating (rather than the marketing) airline because the operating airline is the one that makes pricing decisions. Therefore, we calculate the number of competitors on a route, as well as carrier fixed effects, based on the operating airline.

3.4 Empirical Strategy

3.4.1 Specification

Our empirical strategy seeks to estimate the degree to which competition affects jet fuel pass-through in airfares. We estimate various forms of equation 3.2 below. Our dependent variable ($Airfare_{irt}$) is the weighted (by passengers) average airfare for a given origin-destination (r), airline (i), and month(t).[8] For example, one observation in this specification is the weighted average of airfares for all itineraries that traveled from Sydney to Melbourne on Qantas during January 2010. This means that we aggregate across cabin classes and different itineraries (for example,

[8]Origin-destination pairs are directional, i.e. Sydney–Melbourne is a different route than Melbourne–Sydney.

non-stop itineraries and itineraries that have a stop in Canberra).

$$Airfare_{irt} = \alpha + \beta_1 Comps_{rt} + \beta_2 P^{JF}_{irt} + \beta_3 Comps_{rt} \times P^{JF}_{irt} + \omega Q_t + \delta M_t + \mathbf{X}_{rt} + \epsilon_{irt} \quad (3.2)$$

On the right-hand side, we have the average cost of jet fuel for each origin-destination-airline-month observation (P^{JF}_{irt}), the number of competitors on a route ($Comps_{rt}$), and the interaction of these two terms. Construction of the jet fuel price variable is discussed in Section 3.3. The number of competitors is calculated by counting the number of airlines that have at least 1% market share and fly between the origin and destination, regardless of the number of stops.

We expect β_1 to be negative and β_2 and β_3 to be positive. Our coefficients of interest are β_2 and β_3. The average pass-through rate on a given route can be calculated as $\beta_2 + \beta_3 \times Comps_{rt}$.

3.4.2 Identification

Because airlines typically sign long-term contracts with airports for gate access, and air schedules are set far in advance, we assume that the number of competitors on a route is fixed in the short run. While the number of competitors may be endogenous, it will not vary endogenously with the price of jet fuel – it takes time for an airline to change their route schedule and to enter new markets.

The price of jet fuel is determined in a world market and will not be affected by Australian firms' decisions. However, unobserved macroeconomic demand shocks could make the price of jet fuel endogenous; if demand shocks increase both the demand for air travel and the price of jet fuel, our estimate of the pass-through rate would be biased upwards. To control for unobserved macroeconomic shocks, we include quarter-of-sample (Q_t) fixed effects. Demand for air travel is also highly seasonal (e.g. during the summer and holidays), so we include calendar month (M_t) fixed effects.[9]

Some specifications include a set of state- and route-specific controls ($\mathbf{X}_{rt}$). We construct the geometric mean of the origin and destination's local government area (LGA) populations, densities, and building values.[10] We also use the average of the origin and destination LGA's net migration. We use the geometric mean of origin and destination states' population, wages, business expenditures, mineral exploration, and petroleum exploration.[11] These controls will address

[9] We also include the monthly national unemployment rate as a more fine-grained control in some specifications.

[10] An Australian LGA is roughly equivalent to a US MSA.

[11] The geometric mean is better than the arithmetic mean at capturing variation when locations have very different values. For example, consider how the arithmetic and geometric means differ for: a city of one million people and a town of ten thousand vs. a city of one million and a town of twenty thousand. The arithmetic means will be nearly identical, while the geometric means will be further apart. For airline demand, the latter is likely more informative.

route-specific demand and cost shifts.[12] In most specifications, we also use each flight's capacity factor as a control. This is constructed simply as the total number of passengers on a route divided by the total number of seats (for multi-stop itineraries, each leg's capacity factor is weighted by the miles traveled on that leg). We include origin and destination fixed effects and route fixed effects in some specifications. In the most restrictive specifications with route fixed effects, identification of our parameters of interest comes from shocks to jet fuel prices and entry and exit within routes. This is similar to the identification strategy used by Miller, Osborne and Sheu (2017).

Our analyses are weighted by the number of passengers flying each route. We take this approach because we are primarily interested in the major routes. Flights on small planes that only run once a week are a small part of carbon emissions and may not be broadly representative. Standard errors are clustered at the monthly level.

3.5 Results

Table 3.3 presents our first set of primary results. The second row reports the pass-through rate for a monopoly route, while the third row reports how the pass-through rate varies with additional competitors. Our preferred specification is in the last column, which includes the most restrictive sets of controls and fixed effects. The estimates suggest that pass-through in the presence of a monopoly is roughly 110%, with each additional competitor increasing pass-through by 25%. These estimates are relatively stable across the last four columns.[13]

Our finding of pass-through rates above 100% could be an artifact of how we calculate jet fuel costs. Because passengers are a very small amount of the weight on most jets, their marginal contribution to jet fuel costs is small. To calculate the per person jet fuel cost, we take the total jet fuel cost and spread it across the total number of seats. However, many flights in Australia are not flown at capacity. If a plane flies at half capacity, a 100% pass-through rate implies that the airline is only recovering 50% of the marginal fuel costs. Future work will more fully explore the pass-through implications of unfilled seats.

Table 3.4 presents coefficients when our data have been disaggregated by the number of stops.[14]

We use the arithmetic mean for net migration because net migration can be negative, invalidating the geometric mean.

[12]Some states in Australia experienced a resource boom during our sample period, which is why we control for mineral and petroleum exploration.

[13]Appendix Table C.1 runs a similar set of regressions to those in Table 3.3, but it allows for non-linear responses to competition. The coefficients are similar to those from the linear specification. Coefficients on routes that have five competitors are not in line with the rest of the results. However, there are only 60 observations with 5 competitors. We consider those coefficients anomalous.

[14]This means that we have separate observations for all non-stop, one-stop, two-stop, and three-stop flights for a

These coefficients are very similar to our primary results in Table 3.3. Though our point estimates in the last four columns are centered at values below one, the standard errors are large enough that we cannot reject a pass-through rate of one. Furthermore, at the mean level of competition in our sample – one competitor – the sum of the pass-through and interaction coefficients pushes pass-through above one. This suggests that our results in Table 3.3 are not simply due to how we aggregated the data.

Tables 3.5 and 3.6 report results for only non-stop and one-stop flights, respectively.[15] These are differentiated products and we might expect pass-through to vary between them. Indeed, pass-through on non-stop flights is 67% for a monopoly, with a 33% premium for each additional competitor. For one-stop flights, however, pass-through is 150% with virtually no additional changes due to the number of competitors. Even with one or two competitors, pass-through is lower on non-stop flights. One interpretation of these results is that passengers on non-stop flights are more price-elastic than passengers on one-stop flights. This seems unintuitive since non-stop flights are generally more expensive, but certain models of demand (e.g., logit) imply that elasticity increases with price. Another interpretation is that passengers on one-stop flights are low-value types whose price airlines will distort more in order to respect the high-value non-stop passengers' incentive compatibility constraints. Future structural work on pass-through and price discrimination will look to disentangle these competing explanations.

Table 3.7 includes a jet fuel cost and capacity factor interaction term. Our results show that pass-through is very sensitive to how full planes are. Routes with higher capacity factors have higher pass-through rates.

Finally, Tables 3.8 and 3.9 show how pass-through varies for economy and business class seats. Results suggest that business class pass-through is lower and may be less responsive to competition than economy class pass-through.[16] This matches common intuition about business travelers being less elastic (however, the coefficients are not always statistically different from each other). Future work will more formally discuss the effects of price discrimination through different implements (e.g. non-stop vs. one-stop flights, cabin class, etc.), and their implications for pass-through.

given route-carrier-month.

[15] We drop distance controls from the non-stop regression with route fixed effects as these will be collinear.

[16] In the business class regressions, we define competitors as those carriers who also offer business class seats on the same route.

Table 3.3: Regression Results: Baseline

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	-42.263***	-23.968***	-6.861***	-7.467***	-9.302***	-4.268*
	(2.114)	(1.795)	(1.918)	(1.751)	(1.696)	(2.320)
Jet Fuel Cost	2.038***	3.181***	1.036***	1.139***	1.011***	1.102***
	(0.110)	(0.295)	(0.181)	(0.167)	(0.173)	(0.190)
Jet Fuel * Competitors	1.908***	0.892***	0.288***	0.282***	0.339***	0.251***
	(0.108)	(0.093)	(0.061)	(0.055)	(0.053)	(0.072)
Capacity Factor		-91.101***	43.272***	41.592***	28.067***	57.264***
		(14.229)	(6.631)	(8.746)	(8.716)	(8.726)
Thousand KM, Avg		33.291***	72.203***	59.391***	75.144***	83.478***
		(4.390)	(3.435)	(3.082)	(3.578)	(11.089)
Distance Squared		-6.536***	-1.373**	-1.339**	-2.617***	-12.834***
		(0.326)	(0.573)	(0.521)	(0.493)	(2.108)
Quarter-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.495	0.537	0.768	0.784	0.790	0.833
Observations	112965	112965	112965	112965	102297	102168

Notes: The dependent variable is the average airfare at the month-airline-origin-destination level. For example, one observation is all passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation.

Table 3.4: Regression Results: Observations Separated by Number of Stops

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls
Number of Competitors	-43.306***	-21.297***	-6.720***	-6.208***	-7.687***
	(2.034)	(1.524)	(1.845)	(1.767)	(1.637)
Jet Fuel Cost	1.899***	2.096***	0.956***	0.983***	0.863***
	(0.098)	(0.135)	(0.162)	(0.152)	(0.156)
Jet Fuel * Competitors	1.952***	0.693***	0.252***	0.214***	0.267***
	(0.102)	(0.074)	(0.057)	(0.053)	(0.049)
Capacity Factor		-94.830***	42.812***	43.041***	31.354***
		(7.354)	(6.662)	(8.647)	(8.545)
Thousand KM, Avg		52.711***	83.426***	71.752***	84.395***
		(2.541)	(2.693)	(2.523)	(3.070)
Distance Squared		-5.782***	-3.955***	-3.715***	-4.199***
		(0.255)	(0.168)	(0.145)	(0.156)
Quarter-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes
State/Route Controls	No	No	No	No	Yes
R squared	0.483	0.621	0.759	0.772	0.778
Observations	169394	169394	169394	169394	156573

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-number of stops level. For example, one observation is all non-stop passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation.

Table 3.5: Regression Results: Only Non-Stop Flights

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	-42.538***	-21.324***	-8.621***	-8.847***	-10.889***	-5.824**
	(2.230)	(1.567)	(1.739)	(1.851)	(1.833)	(2.394)
Jet Fuel Cost	1.493***	1.704***	0.480***	0.621***	0.459***	0.674***
	(0.093)	(0.145)	(0.165)	(0.148)	(0.160)	(0.167)
Jet Fuel * Competitors	2.108***	0.824***	0.347***	0.361***	0.424***	0.331***
	(0.117)	(0.078)	(0.063)	(0.065)	(0.064)	(0.080)
Capacity Factor		-87.317***	45.896***	44.476***	33.296***	56.050***
		(7.325)	(6.459)	(8.525)	(8.374)	(8.211)
Thousand KM, Avg		50.714***	69.406***	68.161***	84.977***	
		(2.929)	(3.360)	(3.108)	(3.717)	
Distance Squared		-5.760***	-2.635***	-2.921***	-4.672***	
		(0.285)	(0.522)	(0.485)	(0.423)	
Quarter-Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.469	0.622	0.765	0.771	0.778	0.825
Observations	47739	47739	47739	47739	41495	41459

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-number of stops level. For example, one observation is all non-stop passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Only non-stop flights are included.

Table 3.6: Regression Results: Only One-Stop Flights

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	-107.812***	-36.427***	7.511**	6.263*	6.187**	6.941**
	(7.125)	(3.091)	(3.391)	(3.189)	(2.969)	(3.011)
Jet Fuel Cost	2.052***	2.083***	1.573***	1.741***	1.698***	1.516***
	(0.110)	(0.147)	(0.194)	(0.195)	(0.200)	(0.240)
Jet Fuel * Competitors	1.954***	0.259***	-0.037	0.003	-0.027	-0.031
	(0.117)	(0.064)	(0.077)	(0.069)	(0.064)	(0.074)
Capacity Factor		9.241	183.269***	186.846***	203.139***	204.813***
		(25.291)	(23.743)	(26.837)	(28.363)	(34.890)
Thousand KM, Avg		76.301***	85.328***	71.400***	68.159***	19.573**
		(3.136)	(2.952)	(3.091)	(3.180)	(7.496)
Distance Squared		-5.138***	-4.843***	-4.510***	-4.390***	-4.722***
		(0.234)	(0.234)	(0.212)	(0.203)	(0.832)
Quarter-Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.367	0.527	0.703	0.710	0.711	0.756
Observations	86549	86549	86549	86549	81349	81349

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-number of stops level. For example, one observation is all one-stop passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Only one-stop flights are included.

Table 3.7: Regression Results: Including Capacity Interaction

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	-40.981***	-24.651***	-7.241***	-7.984***	-10.094***	-4.945**
	(1.964)	(1.713)	(1.998)	(1.849)	(1.791)	(2.378)
Jet Fuel Cost	0.205	1.853***	0.651**	0.430	-0.080	0.346
	(0.421)	(0.420)	(0.249)	(0.283)	(0.296)	(0.290)
Jet Fuel * Competitors	1.878***	0.948***	0.300***	0.294***	0.359***	0.269***
	(0.103)	(0.101)	(0.065)	(0.060)	(0.059)	(0.078)
Capacity Factor	-121.185***	-120.823***	32.249***	19.947***	-5.347	33.810***
	(11.562)	(10.405)	(5.837)	(7.431)	(7.623)	(6.735)
Jet Fuel * Capacity	2.737***	1.729***	0.508	0.956**	1.459***	1.002**
	(0.623)	(0.398)	(0.313)	(0.388)	(0.407)	(0.380)
Thousand KM, Avg		32.928***	72.347***	59.457***	75.683***	81.885***
		(4.402)	(3.478)	(3.164)	(3.729)	(10.929)
Distance Squared		-6.125***	-1.329**	-1.262**	-2.553***	-12.273***
		(0.350)	(0.571)	(0.525)	(0.507)	(2.046)
Quarter-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.511	0.539	0.768	0.785	0.791	0.834
Observations	112965	112965	112965	112965	102297	102168

Notes: The dependent variable is the average airfare at the month-airline-origin-destination level. For example, one observation is all passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers on a route.

Table 3.8: Regression Results: Economy Class

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	-41.474***	-23.007***	-6.177***	-7.082***	-8.809***	-3.905*
	(2.066)	(1.717)	(1.846)	(1.666)	(1.618)	(2.137)
Jet Fuel Cost	1.987***	3.405***	1.081***	1.174***	1.078***	1.050***
	(0.108)	(0.305)	(0.204)	(0.184)	(0.190)	(0.207)
Jet Fuel * Competitors	1.954***	0.898***	0.288***	0.284***	0.334***	0.241***
	(0.112)	(0.094)	(0.062)	(0.056)	(0.055)	(0.071)
Capacity Factor		-92.753***	42.355***	42.842***	30.875***	61.161***
		(13.766)	(6.336)	(8.513)	(8.542)	(8.461)
Thousand KM, Avg		29.624***	69.203***	56.693***	71.857***	72.093***
		(4.199)	(3.286)	(2.960)	(3.374)	(10.848)
Distance Squared		-6.632***	-1.740***	-1.641***	-2.941***	-11.240***
		(0.338)	(0.522)	(0.475)	(0.445)	(2.058)
Quarter-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.466	0.510	0.752	0.767	0.772	0.820
Observations	118779	118779	118779	118779	107798	107665

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-cabin class level. For example, one observation is all economy class passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Only economy class fares are included.

Table 3.9: Regression Results: Business Class

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	51.242***	53.108***	17.472*	19.260**	21.794**	7.830
	(9.801)	(11.281)	(9.640)	(9.117)	(9.193)	(8.970)
Jet Fuel Cost	5.154***	3.490***	1.547***	1.939***	1.836***	1.266***
	(0.276)	(0.387)	(0.313)	(0.294)	(0.274)	(0.282)
Jet Fuel * Competitors	-1.001***	-1.030***	-0.376**	-0.253	-0.119	0.171
	(0.277)	(0.286)	(0.186)	(0.178)	(0.184)	(0.189)
Capacity Factor		-61.559	153.183***	173.284***	19.089	37.058
		(46.916)	(43.106)	(33.200)	(32.630)	(35.287)
Thousand KM, Avg		110.075***	132.126***	100.710***	82.358***	747.071***
		(13.642)	(11.117)	(11.492)	(13.327)	(135.109)
Distance Squared		-6.241**	3.153*	1.199	9.324***	-109.966***
		(2.592)	(1.824)	(1.936)	(2.160)	(17.755)
Quarter-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.607	0.628	0.743	0.772	0.792	0.830
Observations	22725	22725	22725	22725	21405	21301

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-cabin class level. For example, one observation is all business class passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Only business class fares are included.

3.6 Conclusion

Our main result is that pass-through rates in the Australian airline industry are very high and increase with the amount of competition and the capacity factor of a route. While it is unusual to have both results simultaneously, they are possible in this industry due to the presence of multi-product firms and heterogeneous consumers, as shown through our simulation study. We also find that product differentiation leads to important heterogeneity in the pass-through rate. The amount of competition on a route has a large effect on pass-through for non-stop flights, but it plays almost no role on one-stop routes. The results generally suggest that a carbon tax on the airline industry would be over-shifted onto consumers.

In this paper, we have presented reduced-form evidence on heterogeneity in pass-through rates across different products and markets. In future work, we will estimate demand for different consumer groups, as well as airline pricing and capacity decisions, to better disentangle the various forces at play. This will allow us to estimate changes in welfare due to changes in jet fuel prices and a carbon tax. Additionally, we will calculate changes in emissions that result from such policies. Overall, the results presented here underscore the importance of taking industry-specific conditions into account when estimating pass-through. The likely effects of various policies will greatly depend on pre-existing market structure; intuition from simpler models will not always carry over

Bibliography

Agrawal, David R, Quinton White, and Jonathan Williams. 2017. "Tax Incidence and Market Power: Evidence from the Airline Industry."

Akee, Randall. 2009. "Checkerboards and Coase: The Effect of Property Institutions on Efficiency in Housing Markets." *Journal of Law and Economics*, 52(2): 395–410.

Akerlof, George A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, 488–500.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton university press.

Armantier, Olivier, and Oliver Richard. 2008. "Domestic Airline Alliances and Consumer Welfare." *The RAND Journal of Economics*, 39(3): 875–904.

Berry, Steven, and Panle Jia. 2010. "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economic Journal: Microeconomics*, 2(3): 1–43.

Berry, Steven T. 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, 889–917.

Biewick, Laura R.H. 2011. "Geodatabase of Wyoming Statewide Oil and Gas Drilling Activity to 2010." *U.S. Geological Survey Data Series 625*.

Bleakley, Hoyt, and Joseph Ferrie. 2014. "Land Openings on the Georgia Frontier and the Coase Theorem in the Short- and Long-Run." http://www-personal.umich.edu/~hoytb/Bleakley_Ferrie_Farmsize.pdf.

Borenstein, Severin. 1989. "Hubs and High Fares: Dominance and Market Power in the US Airline Industry." *The RAND Journal of Economics*, 344–365.

Borenstein, Severin, and Nancy L Rose. 1994. "Competition and Price Dispersion in the US Airline Industry." *Journal of Political Economy*, 102(4): 653–683.

Bureau of Land Management. 1983. *The Federal Simultaneous Oil and Gas Leasing System: Important Information about the "Government Oil and Gas Lottery"*. Bureau of Land Management.

Cabral, Marika, Michael Geruso, and Neale Mahoney. 2014. "Does Privatized Health Insurance Benefit Patients or Producers? Evidence from Medicare Advantage."

Callaway, Duncan, Meredith Fowlie, and Gavin McCormick. 2015. "Location, Location, Location: The Variable Value of Renewable Energy and Demand-Side Efficiency Resources." *Energy Institute at Haas–Working paper*, 264.

Chandra, Ambarish, and Mara Lederman. 2015. "Revisiting the Relationship between Competition and Price Discrimination in the Airline Industry."

Cicala, Steve. 2014. "When does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation." *The American Economic Review*, 105(1): 411–444.

Ciliberto, Federico, and Elie Tamer. 2009. "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica*, 77(6): 1791–1828.

Coase, Ronald H. 1960. "The Problem of Social Cost." *The Journal of Law and Economics*, 3: 1–44.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of econometrics*, 92(1): 1–45.

Cullen, Joseph A. 2013. "Measuring the Environmental Benefits of Wind-Generated Electricity." *American Economic Journal: Economic Policy*, 5(4): 107–133.

Cullen, Joseph A, and Erin T Mansur. 2014. "Inferring Carbon Abatement Costs in Electricity Markets: A Revealed Preference Approach using the Shale Revolution."

Dang, Tri Vi, Gary Gorton, Bengt Holmström, and Guillermo Ordonez. 2017. "Banks as Secret Keepers." *The American Economic Review*, 107(4): 1005–1029.

Davis, Lucas W, and Catherine Hausman. 2014. "The Value of Transmission in Electricity Markets: Evidence from a Nuclear Power Plant Closure."

Davis, Lucas W, and Lutz Kilian. 2011. "The Allocative Cost of Price Ceilings in the US Residential Market for Natural Gas." *Journal of Political Economy*, 119(2): 212–241.

Dorsey-Palmateer, Reid. 2014. "Effects of Wind Power Intermittency on Generation and Emissions." Working Paper.

EIA. 2001. "Impact of Interruptible Natural Gas Service on Northeast Heating Oil Demand." http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/interruptible_service_northeast/interrupt.pdf.

EIA. 2012. "Price Ratio of Crude Oil to Natural Gas Continues to Increase." http://www.eia.gov/todayinenergy/detail.cfm?id=5830.

Fabra, Natalia, and Mar Reguant. 2014. "Pass-Through of Emissions Costs in Electricity Markets." *The American Economic Review*, 104(9): 2872–2899.

Fairfax, Sally K, and Carolyn E Yale. 1987. *Federal Lands: A Guide to Planning, Management, and State Revenues.* Island Press.

Fowlie, Meredith, Mar Reguant, and Stephen P Ryan. 2016. "Market-based Emissions Regulation and Industry Dynamics." *Journal of Political Economy*, 124(1): 249–302.

Ganapati, Sharat, Joseph S Shapiro, and Reed Walker. 2016. "Energy Prices, Pass-Through, and Incidence in US Manufacturing."

Gerardi, Kristopher S, and Adam Hale Shapiro. 2009. "Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry." *Journal of Political Economy*, 117(1): 1–37.

Goldberg, Pinelopi Koujianou, and Rebecca Hellerstein. 2013. "A Structural Approach to Identifying the Sources of Local Currency Price Stability." *The Review of Economic Studies*, 80(1): 175–210.

Goolsbee, Austan, and Chad Syverson. 2008. "How do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines." *The Quarterly Journal of Economics*, 123(4): 1611–1633.

Graff Zivin, Joshua S, Matthew J Kotchen, and Erin T Mansur. 2014. "Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and other Electricity-Shifting Policies." *Journal of Economic Behavior & Organization*, 107: 248–268.

Greene, William H. 2012. *Econometric Analysis, 7th.*

Hausman, Catherine, and Ryan Kellogg. 2015. "Welfare and Distributional Implications of Shale Gas." National Bureau of Economic Research Working Paper 21115.

Heath, Garvin A, Patrick O'Donoughue, Douglas J Arent, and Morgan Bazilian. 2014. "Harmonization of Initial Estimates of Shale Gas Life Cycle Greenhouse Gas Emissions for Electric Power Generation." *Proceedings of the National Academy of Sciences*, 111(31): E3167–E3176.

Hendricks, Kenneth, and Robert H. Porter. 1988. "An Empirical Study of an Auction with Asymmetric Information." *The American Economic Review*, 78: 865–863.

Holladay, J Scott, and Jacob LaRiviere. 2014. "The Impact of Cheap Natural Gas on Marginal Emissions from Electricity Generation and Implications for Energy Policy." http://web.utk.edu/~jlarivi1/docs/HL_gas.pdf.

Interagency Working Group on Social Cost of Carbon, United States Government. 2013. "Technical Update on the Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866."

Ito, Koichiro. 2014. “Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing.” *American Economic Review*, 104(2): 537–563.

Kaffine, Daniel T, Brannin J McBee, and Jozef Lieskovsky. 2013. “Emissions Savings from Wind Power Generation in Texas.” *The Energy Journal*, 34(1): 155–175.

Kilian, Lutz. 2015. “energy price shocks.” In *The New Palgrave Dictionary of Economics*. , ed. Steven N. Durlauf and Lawrence E. Blume, 1–15. Palgrave Macmillan.

Lafrancois, Becky A. 2012. “A Lot Left Over: Reducing CO2 Emissions in the United States Electric Power Sector Through the use of Natural Gas.” *Energy Policy*, 50: 428–435.

Lewis, Eric. 2015. “Federal Environmental Protection and the Distorted Search for Oil and Gas.”

Linn, Joshua, Erin Mastrangelo, and Dallas Burtraw. 2014. “Regulating Greenhouse Gases from Coal Power Plants under the Clean Air Act.” *Journal of the Association of Environmental and Resource Economists*, 1(1): 97–134.

Linn, Joshua, Lucija Muehlenbachs, and Yushuang Wang. 2014. “How Do Natural Gas Prices Affect Electricity Consumers and the Environment?”

Li, Shanjun. 2014. “Better Lucky Than Rich? Welfare Analysis of Automobile License Allocations in Beijing and Shanghai.” *Welfare Analysis of Automobile License Allocations in Beijing and Shanghai (March 1, 2014)*.

Lu, Xi, Jackson Salovaara, and Michael B. McElroy. 2012. “Implications of the Recent Reductions in Natural Gas Prices for Emissions of CO2 from the US Power Sector.” *Environmental Science & Technology*, 46(5): 3014–3021.

Miller, Nathan H, Matthew Osborne, and Gloria Sheu. 2017. “Pass-Through in a Concentrated Industry: Empirical Evidence and Regulatory Implications.” *The RAND Journal of Economics*, 48(1): 69–93.

Myerson, Roger B. 1985. “Analysis of Two Bargaining Problems with Incomplete Mechanisms.” In *Game-theoretic models of bargaining*. , ed. Alvin E Roth, Chapter 7, 115–147. Cambridge University Press.

Myerson, Roger B., and Mark A. Satterthwaite. 1983. “Efficient Mechanisms for Bilateral Trading.” *Journal of Economic Theory*, 29: 265–281.

Novan, Kevin. 2014. “Valuing the Wind: Renewable Energy Policies and Air Pollution Avoided.” *American Economic Journal: Economic Policy (forthcoming)*.

Ramberg, David J, and John E Parsons. 2012. “The Weak Tie between Natural Gas and Oil Prices.” *The Energy Journal*, 33(2).

Ryan, Stephen P. 2012. “The Costs of Environmental Regulation in a Concentrated Industry.” *Econometrica*, 80(3): 1019–1061.

Samuelson, William. 1984. "Bargaining under Asymmetric Information." *Econometrica: Journal of the Econometric Society*, 995–1005.

Samuelson, William. 1985. "A Comment on the Coase Theorem." In *Game-theoretic models of bargaining*. , ed. Alvin E Roth, Chapter 15, 321–339. Cambridge University Press.

Venkatesh, Aranya, Paulina Jaramillo, W Michael Griffin, and H Scott Matthews. 2012. "Implications of Changing Natural Gas Prices in the United States Electricity Sector for SO2, NOX and Life Cycle GHG Emissions." *Environmental Research Letters*, 7(3): 034018.

Villar, Jose A, and Frederick L Joutz. 2006. "The Relationship between Crude Oil and Natural Gas Prices." *Energy Information Administration, Office of Oil and Gas*, 1–43.

Weyl, E Glen, and Michael Fabinger. 2013. "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition." *Journal of Political Economy*, 121(3): 528–583.

Wiggins, Steven N, and Gary D Libecap. 1985. "Oil field unitization: Contractual Failure in the Presence of Imperfect Information." *The American Economic Review*, 368–385.

APPENDIX A

First Chapter Appendices

A.1 Data Sources and Identification of Firms and Individuals

Records on lotteries were scanned from paper records at the BLM office in Laramie Wyoming. These records include first information on parcels that would be offered in the lottery, with an example from June 1975 in Figure A.1. These records also contain information on the first-, second-, and third-place winners as well as the total number of entries. The first place winner has information both on name as well as address; the second- and third-place winners only have names.

This data is publicly available from the Wyoming BLM. The Michigan IRB panel ruled that this data is not regulated.

Data were double-blind entered using a data digitization service and are guaranteed to be 99.95% accurate.

We identify firms from whether words such as “Co.”, “Corp. ”, “Corporation”, “Co.”, “Inc. ”, “Ltd.”, “Limited ”, “Associates”, “Oil”, “Gas”, and “Industries”appear in the name of the winner. We also include as firms those that are obviously firms but not easily categorized from this rule (for example, “Michigan Wisconsin Pipe Line”).

We also explicitly list first-place individuals as individuals rather than firms even if their address information suggests that they are associated with a firm (e.g., John Doe, Acme Co., Acme Wyoming 80000) We do this for two reasons. First, if these individuals had appeared as second- or third-place winners, we would not observe the address/firm information, and we would categorize them as individuals. Second, we cannot determine whether these individuals were entering the lottery on behalf of the firm or merely using the firm as a personal address. To the extent that this is an issue, our analysis will misclassify individuals as firms. We reviewed the winners in our sample and did not find evidence that individuals entering on behalf of firms was widespread. Because our dependent variable is binary, any such errors will attenuate our coefficients towards zero.

A.2 Additional Reassignment Analysis

This section presents additional reassignment analysis. Table A.1 presents regression results using Equation 1.1 to analyze the time until initial lease reassignment. Table A.2 demonstrates that reassignment generally preceded drilling.

Table A.1: Regression Results: Lease Transfers, Restricted Sample

	(1) Transfer Within First Year	(2) Transfer During Years 2-10	(3) Transfer After Ten Years	(4) Never Transfer
Firm in first place	-0.223***	-0.005	0.008*	0.220***
	(0.021)	(0.030)	(0.005)	(0.028)
Nearby Production Flag	0.027	-0.003	0.004	-0.029
	(0.038)	(0.031)	(0.007)	(0.035)
Firm/Nearby Prod Interaction	-0.040	-0.031	0.012	0.056
	(0.053)	(0.063)	(0.017)	(0.069)
Offers & Acreage Controls	Yes	Yes	Yes	Yes
Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.157	0.040	0.047	0.145
Observations	1692	1692	1692	1692

Notes: Regressions for restricted sample where there was exactly one firm among the three winners. The first column dependent variable is whether the lease was transferred within 0-1 years of the start. The second column variable is whether it was transferred for the first time within 2-10 years. The third column is whether it was transferred for the first time more than 10 years after start date. And the fourth column is an indicator for whether it never was transferred. For each of these specifications, leases that do not appear in the LR2000 are omitted.

Figure A.1: Parcel Offering Examples

SIXTH PRINCIPAL MERIDIAN

WYOMING

#1096 W 0316078
T 17 N, R 60 W, Laramie
Sec 6: Lots 3, 7, S½NE¼, SE¼NW¼, E½SW¼, N½SE¼
8: N½NW¼
T 18 N, R 60 W
Sec 29: NE¼NW¼, W½SW¼
30: Lot 1, NW¼NE¼, NE¼NW¼, NE¼SE¼
31: NE¼, E½SW¼, W½SE¼
32: S½NE¼, W½
1444.21 A

#1097 W 23542
T 26 N, R 60 W, Goshen
Sec 3: SW¼NW¼, NW¼SW¼
4: SE¼NE¼, E½SW¼, NE¼SE¼
8: SE¼NE¼, SE¼NW¼
10: Lot 4
15: Lot 1, SE¼NW¼, NE¼SW¼
22: Lot 3, SE¼NW¼, NE¼SW¼
28: NE¼SE¼
29: E½SW¼
32: SW¼, S½SE¼
33: SW¼NE¼, NW¼SE¼
1099.17 A

#1098 W 5212
T 35 N, R 60 W, Niobrara
Sec 18: NE¼NE¼
40.00 A

#1099 W 43988
T 36 N, R 60 W, Niobrara
Sec 31: Lot 1, NE¼, E½NW¼, NE¼SE¼
32: S½
638.78 A

#1100 W 0309148
T 36 N, R 62 W, Niobrara
Sec 13: W½NE¼, E½NW¼
160.00 A

#1101 W 0316090
T 39 N, R 62 W, Niobrara
Sec 35: NW¼
160.00 A

#1102 W 0310335
T 43 N, R 62 W, Weston
Sec 11: W½SW¼, W½SE¼
14: Lots 1, 4, E½NW¼, W½SE¼
394.53 A

#1106 W 5223
T 38 N, R 63 W, Niobrara
Sec 13: S½NE¼, SE¼
14: NW¼NE¼, S½NE¼, NW¼, SE¼
23: NW¼NE¼
720.00 A

#1107 W 0316100
T 43 N, R 63 W, Weston
Sec 29: SE¼SE¼
40.00 A

#1108 W 0324155
T 37 N, R 64 W, Niobrara
Sec 13: N½SW¼
80.00 A

#1109 W 23556
T 40 N, R 64 W, Niobrara
Sec 7: Lot 3, S½NE¼, NE¼SW¼
166.31 A

#1110 W 0220694-A
T 40 N, R 64 W, Niobrara
Sec 13: SW¼
25: SW¼NW¼, E½SE¼
280.00 A

#1111 W 0220694
T 40 N, R 64 W, Niobrara
Sec 14: W½NW¼
80.00 A

#1112 W 39112
T 41 N, R 64 W, Weston
Sec 4: S½SE¼
14: S½SE¼
160.00 A

#1113 W 0314722
T 46 N, R 64 W, Weston
Sec 17: NW¼NW¼
40.00 A

#1114 W 11816-A
T 48 N, R 64 W, Weston
Sec 4: SE¼SW¼, NE¼SE¼, S½SE¼
160.00 A

#1115 W 0316107
T 35 N, R 65 W, Niobrara
Sec 9: E½SE¼
10: SW¼NW¼, W½SW¼
200.00 A

Notes: Raw data example of parcels offered in the June 1975 lottery.

Table A.2: Regression Results: Drilling Before Reassignment

	Drilling	Production	P | D
Firm Winner	0.008	0.007	0.181
	(0.008)	(0.005)	(0.224)
Nearby Production Flag (2.6 Miles)	0.039***	0.018**	0.010
	(0.013)	(0.008)	(0.200)
Firm/Nearby Prod Interaction	0.002	0.015	0.115
	(0.026)	(0.017)	(0.246)
intercept	0.010***	0.005**	0.444**
	(0.003)	(0.002)	(0.176)
R squared	0.013	0.011	0.063
Observations	1692	1692	36

A: All parcels, including never reassigned, restricted sample

	Drilling	Production	P | D
Firm Winner	0.011	0.007	0.183
	(0.007)	(0.005)	(0.141)
Nearby Production Flag (2.6 Miles)	0.026***	0.012***	0.122
	(0.005)	(0.004)	(0.085)
Firm/Nearby Prod Interaction	0.005	0.017	0.092
	(0.021)	(0.014)	(0.305)
intercept	0.004	-0.011***	-0.362***
	(0.011)	(0.002)	(0.117)
R squared	0.025	0.019	0.405
Observations	10191	10191	184

B: All parcels, including never reassigned, full sample

Notes: This table examines the probability of drilling and production activity prior to initial reassignment. Panel A uses our restricted sample, while panel B uses the full sample. Both panels are limited to leases with LR2000 data.

A.3 Full Sample Analogues to Primary Results

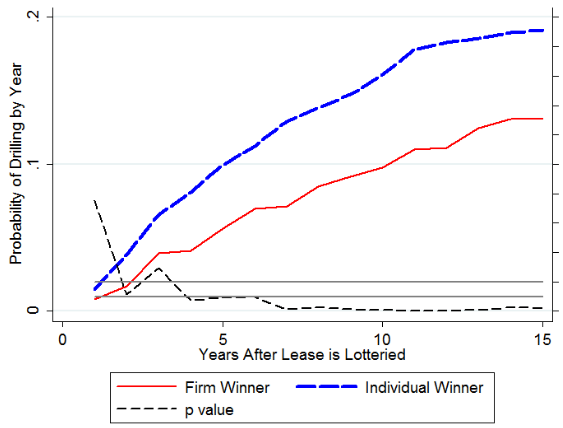
This section presents the full sample analogues to our primary results about nearby production, which use our restricted sample. Figure A.2 is analogous to Figure 1.4, Table A.3 is analogous to Table 1.5, and Figure A.3 is analogous to Figure 1.5.

Table A.3: Regression Results: Main Results, Full Sample

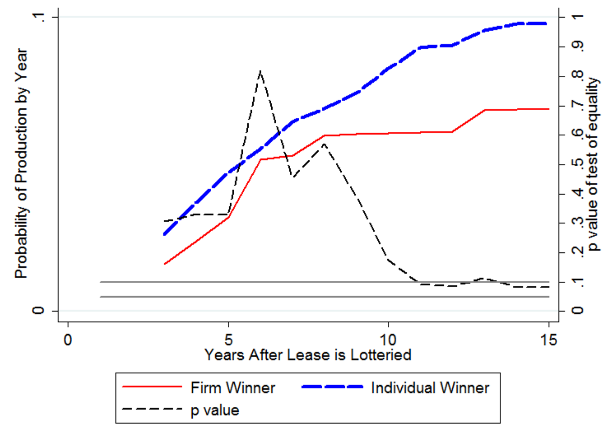
	(1) Reassign Probability	(2) Log Time to Reassign	(3) Drill	(4) Prod	(5) P | D
Nearby Production Flag	0.017	-0.070	0.111***	0.058***	0.079*
	(0.014)	(0.073)	(0.026)	(0.011)	(0.046)
Firm Winner	-0.211***	0.875***	0.011	-0.000	-0.074
	(0.026)	(0.086)	(0.011)	(0.005)	(0.069)
Firm/Nearby Prod Interaction	-0.056	0.219	-0.083***	-0.029	0.092
	(0.056)	(0.176)	(0.026)	(0.018)	(0.116)
Offers & Acreage Controls	Yes	Yes	Yes	Yes	Yes
Lottery Fixed Effects	Yes	Yes	Yes	Yes	Yes
R squared	0.130	0.123	0.078	0.061	0.151
Observations	10762	8097	10762	10762	904

Notes: This table uses the probability of reassignment by twelve years (1), length of time given reassignment (2), probability of drilling by twelve years (3), probability of production by twelve years (4), and probability of production given drilling by twelve years (5) as dependent variables. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on. This table uses our full sample. Column (2) does not correct for selection into reassignment. Point estimates using a Heckman two-step are similar.

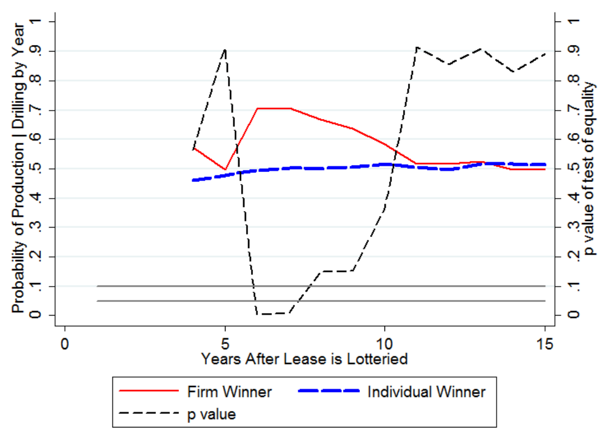
Figure A.2: Full Sample Results: Leases with Nearby Production



(a) Drilling



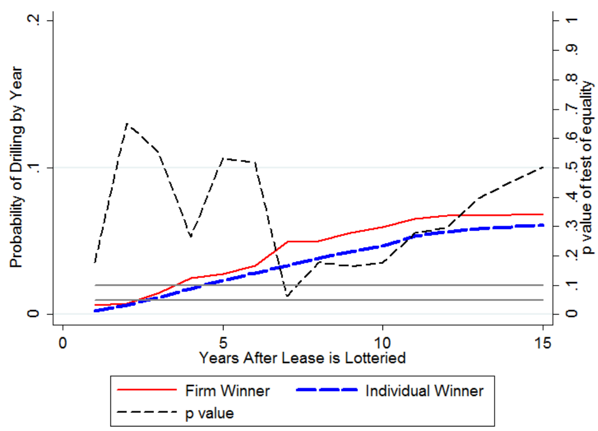
(b) Production



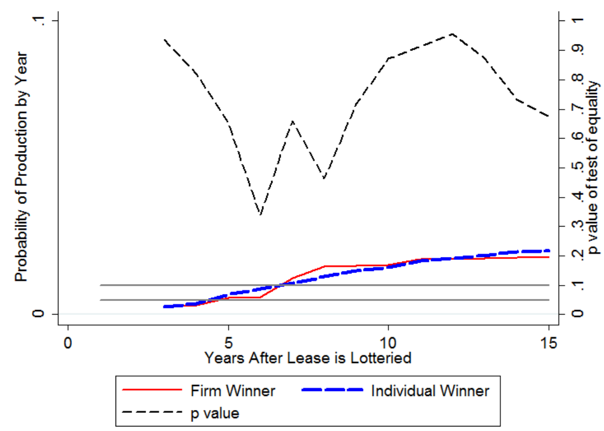
(c) Production Given Drilling

Notes: Full sample using observables to control for endogenous entry. Estimates are reported for parcels with nearby production. The firm effect is the sum of the coefficient on an indicator for whether the first place winner was a firm, plus the coefficient on an interaction between the firm winner indicator and an indicator for nearby production. The right vertical axis gives the p value of a test that the two means are not equal.

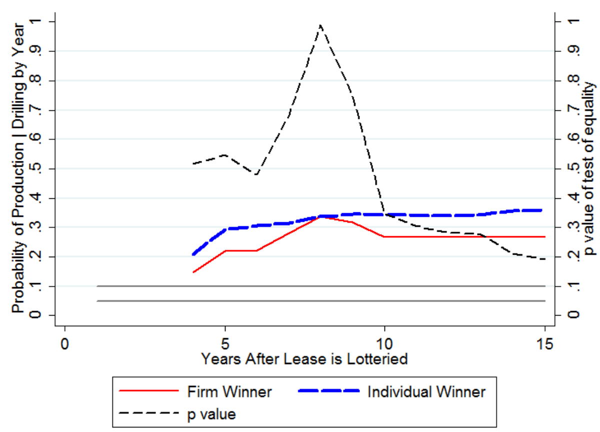
Figure A.3: Full Sample Results: Leases without Nearby Production



(a) Drilling



(b) Production



(c) Production Given Drilling

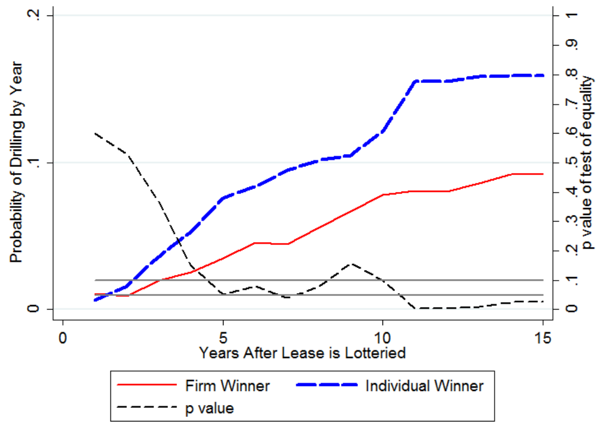
Notes: Full sample using observables to control for endogenous entry. Estimates are reported for parcels without nearby production. The firm effect is the coefficient on an indicator for whether the first place winner was a firm. The right vertical axis gives the p value of a test that the two means are not equal.

A.4 Alternative Nearby Production Specifications

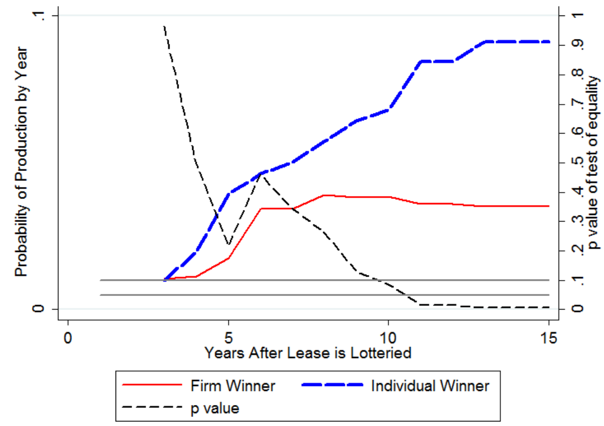
In this appendix we consider a variety of alternative nearby production specifications. Figures A.4 and A.5 define nearby production as occurring over the last 10 years. Figures A.6 and A.7 define nearby production as occurring within 1.7 miles. Figures A.8 and A.9 define nearby production as occurring within 3.2 miles. Figures A.10 and A.11 use an alternate data source that includes productive gas wells and defines recent by the first production year, not the drilling year. Figures A.12 and A.13 repeat the exercise in Figures A.10 and A.11, but define nearby production as occurring within 3.2 miles. Our central results about nearby production are broadly consistent over this range of alternative specifications.

Figures A.14 and A.15 look at nearby drilling instead of nearby production. We limit "nearby" to only include the closest 1.7 miles as well over half of leases have had unsuccessful wells within 2.6 miles. The results are qualitatively similar to our main results, though the magnitudes are smaller and not always statistically significant. This finding is unsurprising. Unproductive wells will have two important differences from productive ones. First, the owner of an unproductive well will be less eager to purchase nearby leases – they have already determined that the area is dry. Second, the information asymmetry from an unproductive well is likely not as large. Both parties will be able to agree that the nearby firm has not found oil – considerably easier than agreeing about the quantity and quality of production from a nearby producing well. As the quality of information causing information asymmetry decreases, the negative effects from information asymmetry should dissipate.

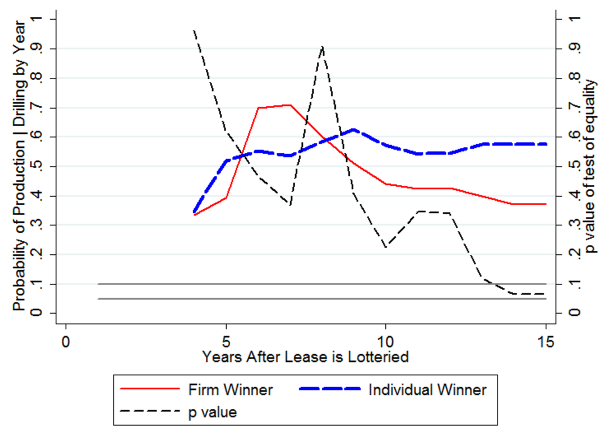
Figure A.4: Restricted Sample Results: Longer Nearby Production Window



(a) Drilling



(b) Production

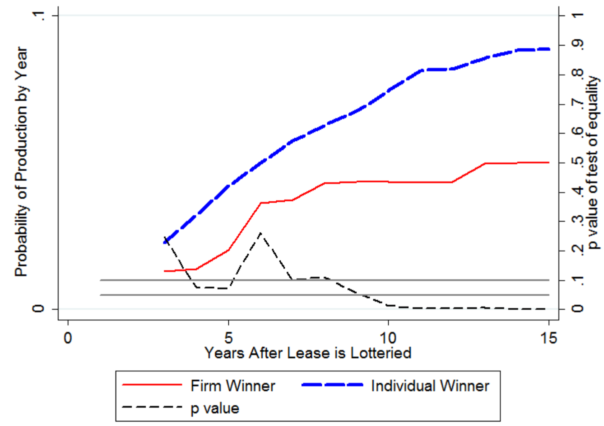


(c) Production Given Drilling

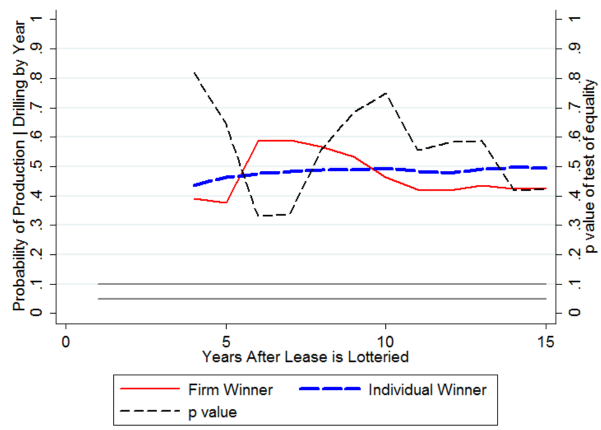
Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by allowing nearby production to occur over the 10 years preceding a lease.

Figure A.5: Full Sample Results: Longer Nearby Production Window

(a) Drilling



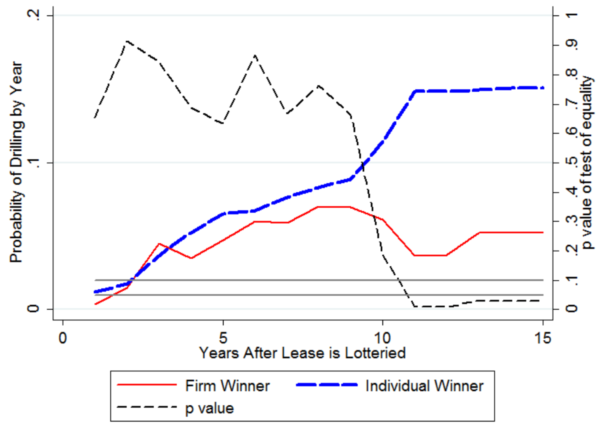
(b) Production



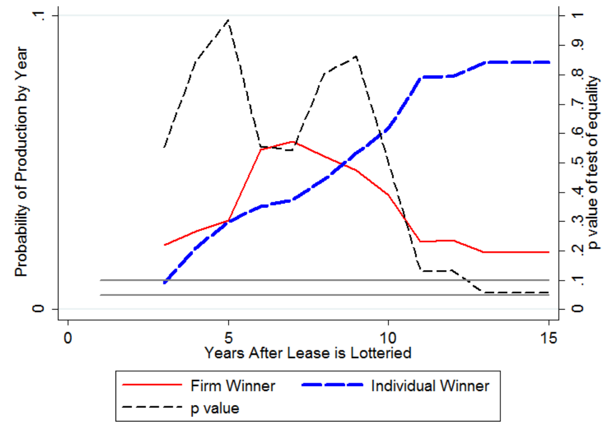
(c) Production Given Drilling

Notes: Full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. This differs from our primary specification by allowing nearby production to occur over the 10 years preceding a lease.

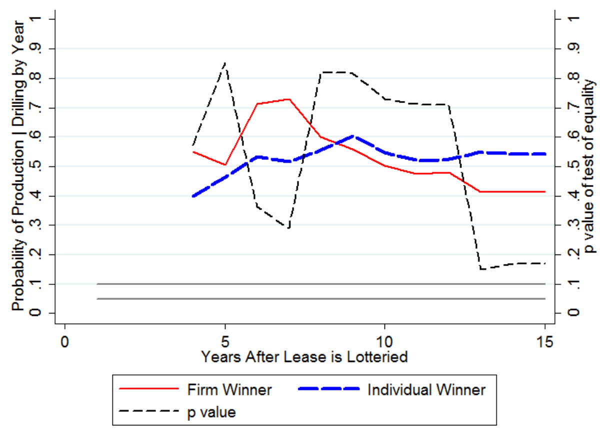
Figure A.6: Restricted Sample Results: Tighter Nearby Definition



(a) Drilling



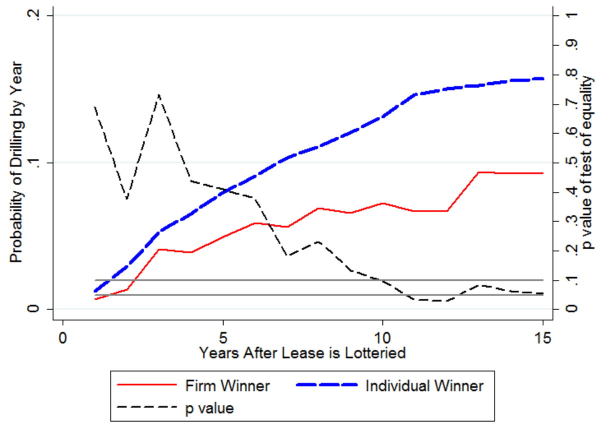
(b) Production



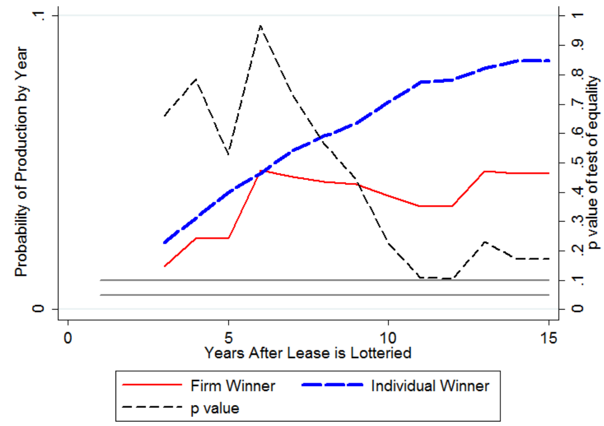
(c) Production Given Drilling

Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by defining nearby production as occurring within 1.7 miles of a lease.

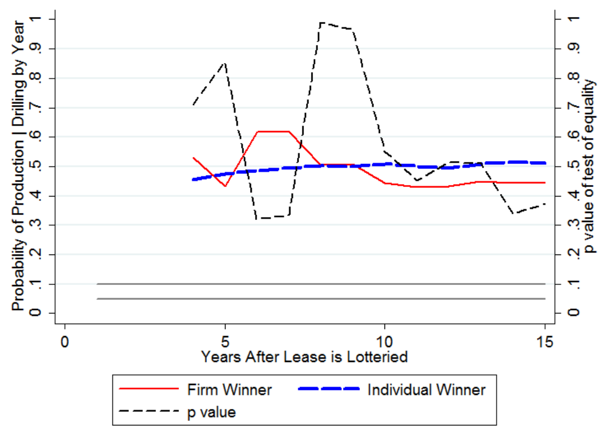
Figure A.7: Full Sample Results: Tighter Nearby Definition



(a) Drilling



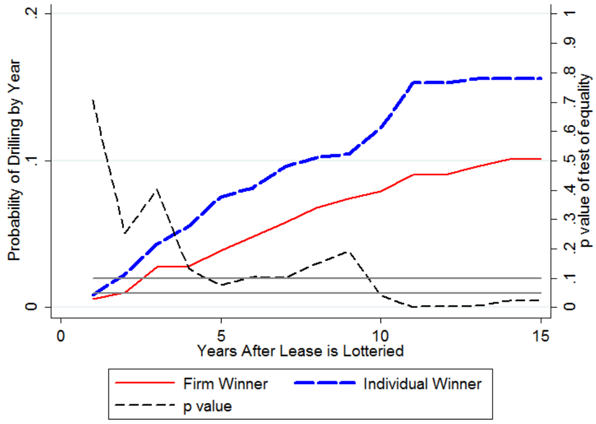
(b) Production



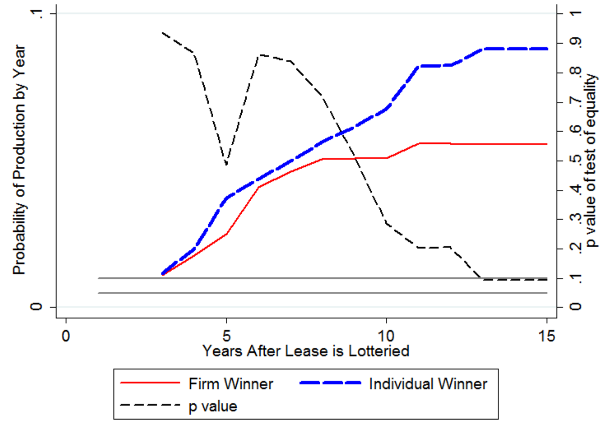
(c) Production Given Drilling

Notes: Full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. This differs from our primary specification by defining nearby production as occurring within 1.7 miles of a lease.

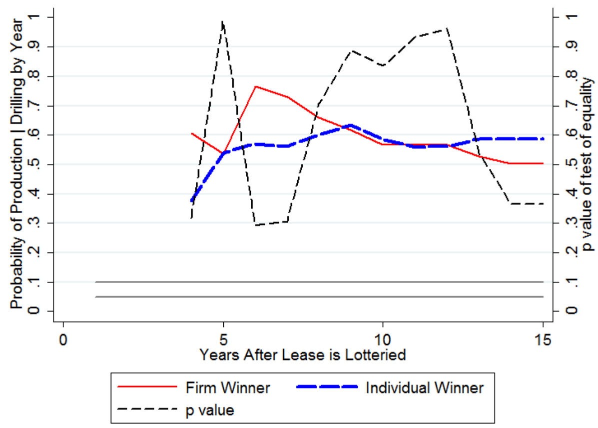
Figure A.8: Restricted Sample Results: Looser Nearby Definition



(a) Drilling



(b) Production

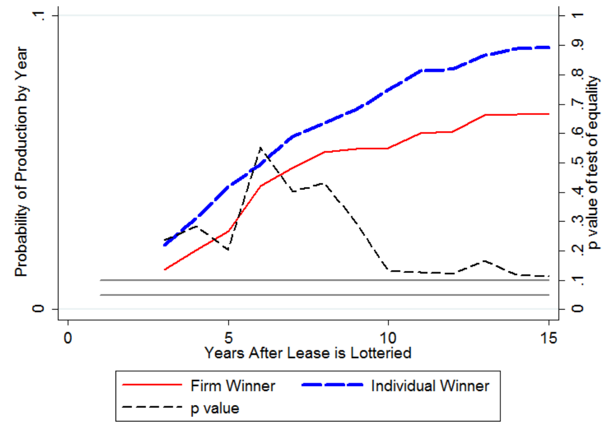


(c) Production Given Drilling

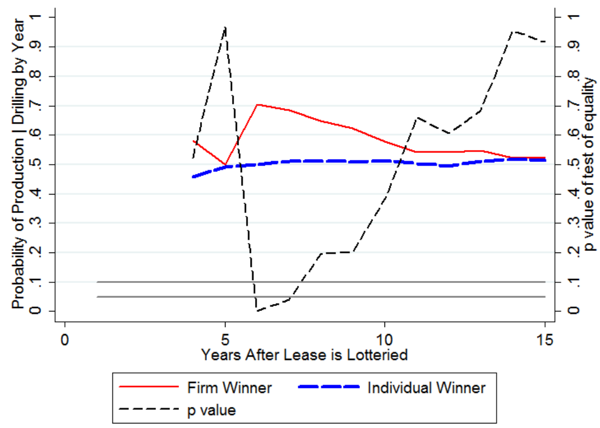
Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by defining nearby production as occurring within 3.2 miles of a lease.

Figure A.9: Full Sample Results: Looser Nearby Definition

(a) Drilling



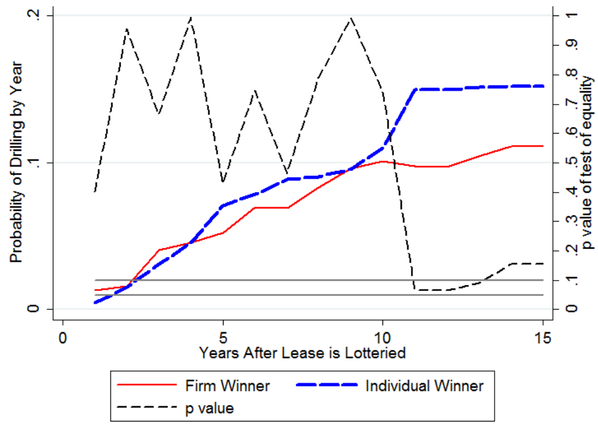
(b) Production



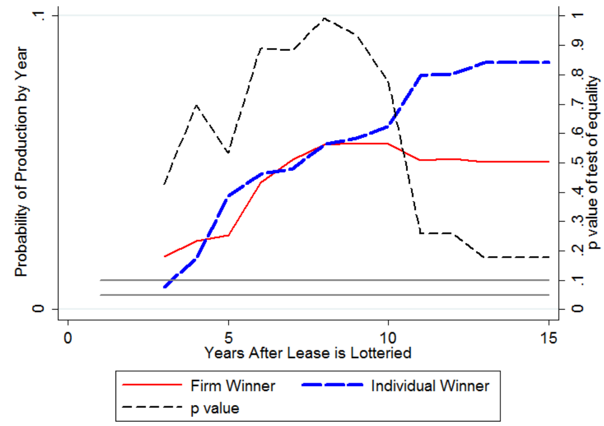
(c) Production Given Drilling

Notes: Full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. This differs from our primary specification by defining nearby production as occurring within 3.2 miles of a lease.

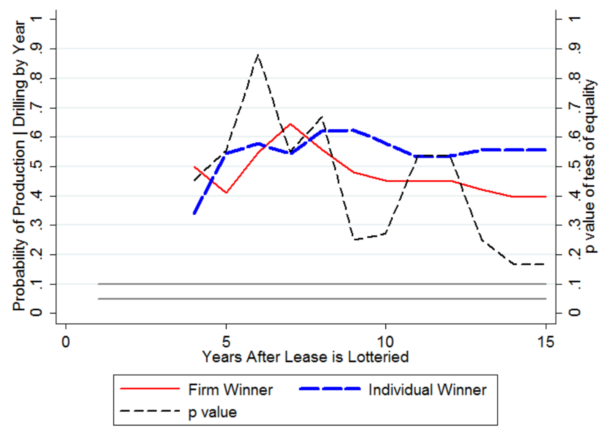
Figure A.10: Restricted Sample Results: Including Gas Production



(a) Drilling



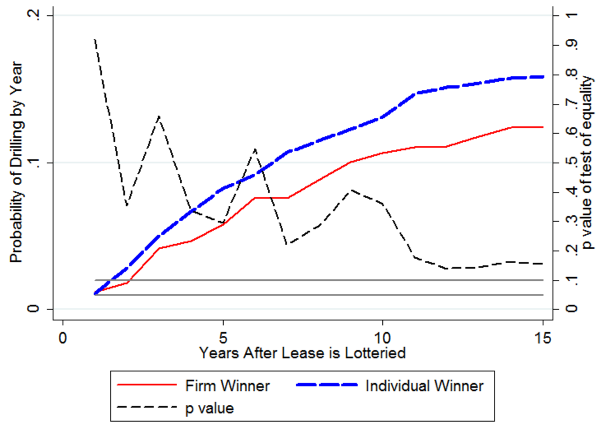
(b) Production



(c) Production Given Drilling

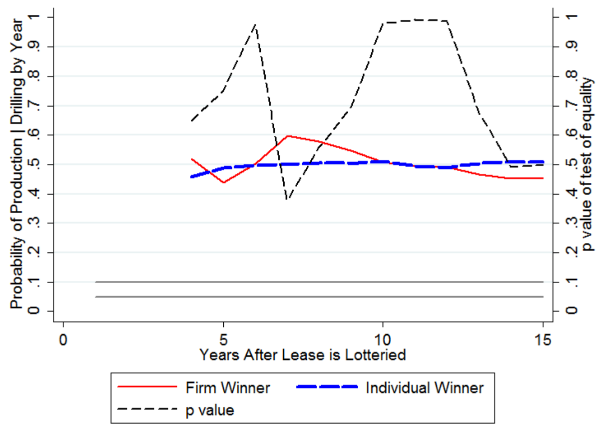
Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by using an alternate data source that includes productive gas wells and defines recent by the first production year, not the drilling year.

Figure A.11: Full Sample Results: Including Gas Production



(a) Drilling

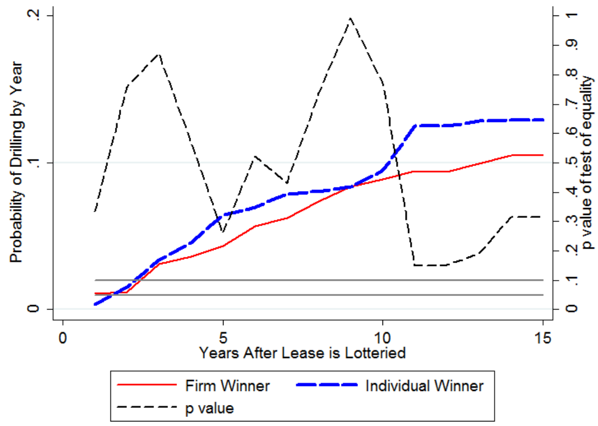
(b) Production



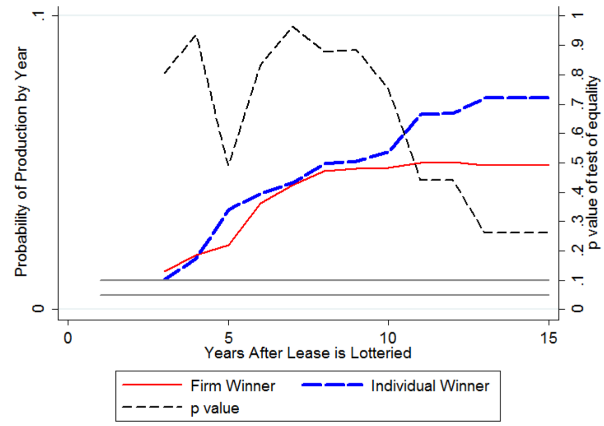
(c) Production Given Drilling

Notes: Full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. This differs from our primary specification by using an alternate data source that includes productive gas wells and defines recent by the first production year, not the drilling year.

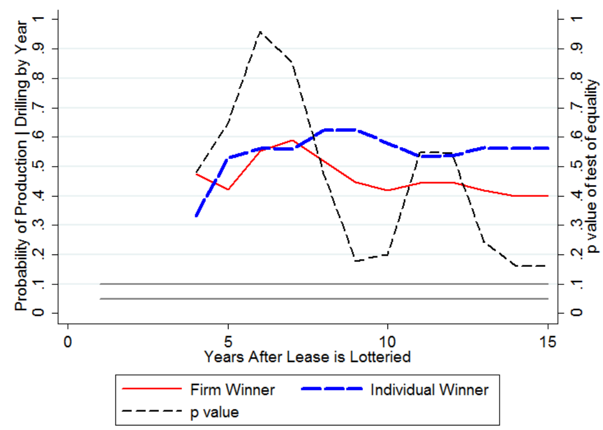
Figure A.12: Restricted Sample Results: Including Gas Production & Looser Nearby Definition



(a) Drilling



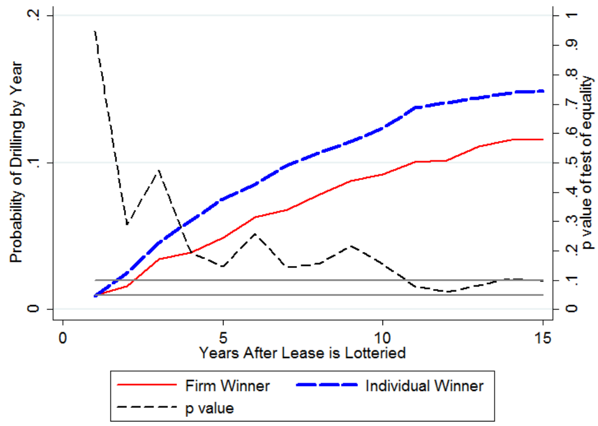
(b) Production



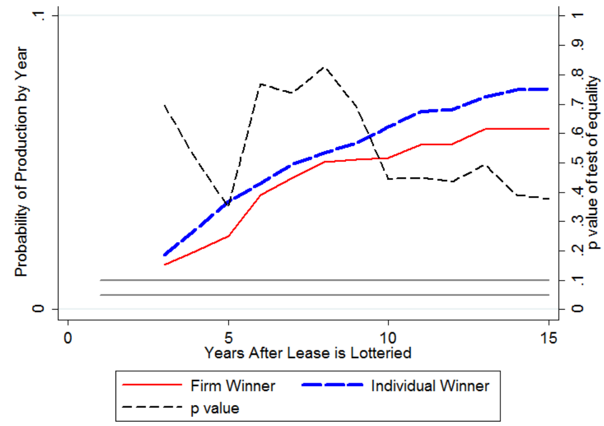
(c) Production Given Drilling

Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by using an alternate data source that includes productive gas wells and defines recent by the first production year, not the drilling year, and defining nearby as within 3.2 miles.

Figure A.13: Full Sample Results: Including Gas Production & Looser Nearby Definition



(a) Drilling

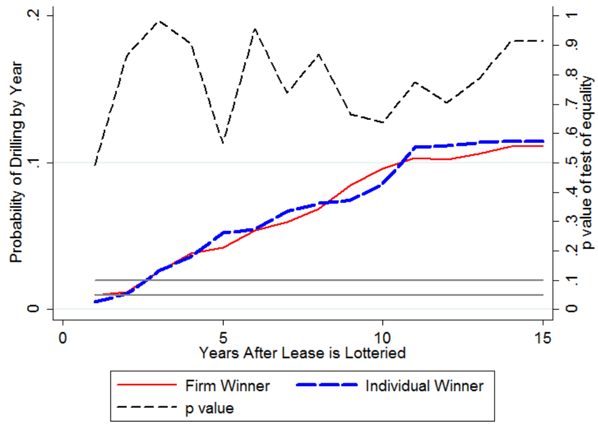


(b) Production

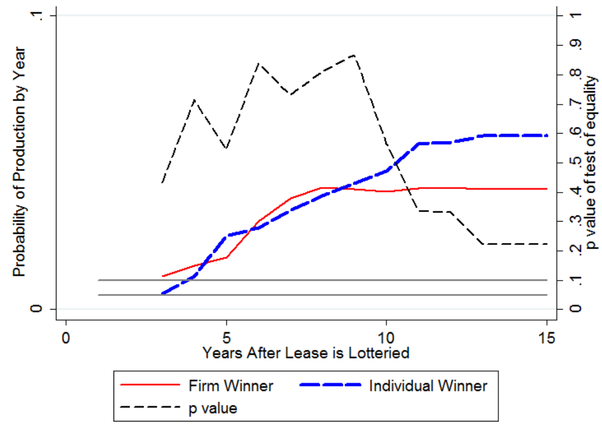
(c) Production Given Drilling

Notes: Full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. This differs from our primary specification by using an alternate data source that includes productive gas wells and defines recent by the first production year, not the drilling year, and defining nearby as within 3.2 miles.

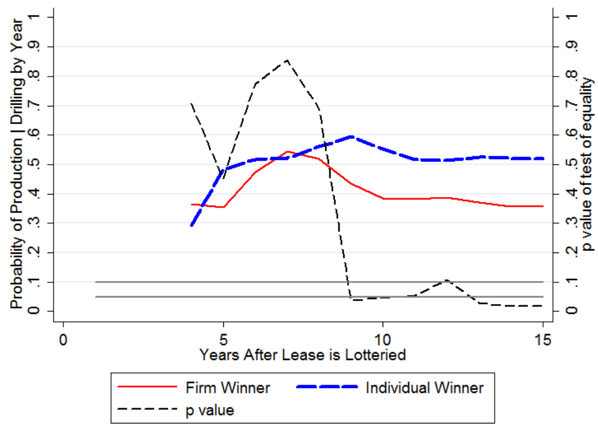
Figure A.14: Restricted Sample Results: Including Gas Production & Tighter Nearby Definition



(a) Drilling



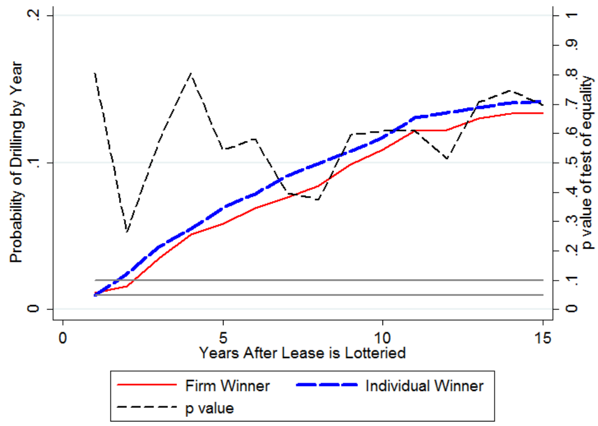
(b) Production



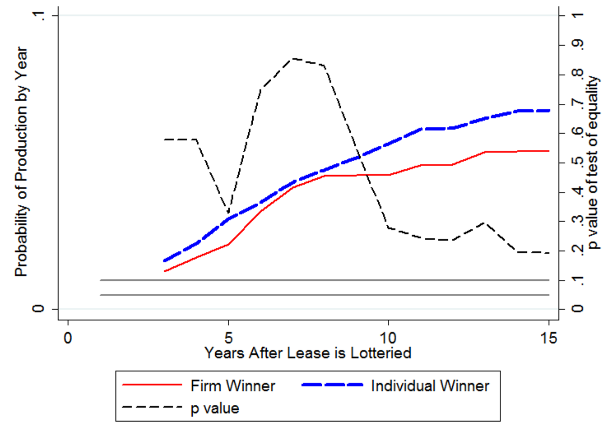
(c) Production Given Drilling

Notes: Restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by defining nearby production as all wells drilled within 1.7 miles.

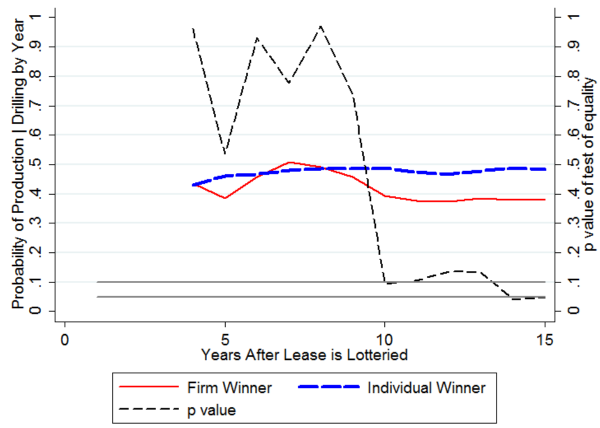
Figure A.15: Full Sample Results: Including Gas Production & Tighter Nearby Definition



(a) Drilling



(b) Production



(c) Production Given Drilling

Notes: Full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. This differs from our primary specification by defining nearby production as all wells drilled within 1.7 miles.

A.5 Drilling Durations and Production Quantities

This appendix includes regressions that look at drilling durations and production quantities. Because many drilling rigs are rented out by the day, the number of days spent drilling is a rough proxy for drilling costs. Production quantities are a rough proxy for revenue. Due to the highly variable nature of these measures, our results are not precise. Table A.4 presents the results of this analysis.

The first row shows that drilling time (Columns 1 & 2) and production quantities (Columns 3 & 4) do not have statistically significant differences depending on the initial lease winner's identity. The point estimates suggest firms had lower average production, but they are insignificant; we do not interpret them as meaningful. We do not have a reason to expect that these types of wells would have *lower* quantities produced.[1]

The third row looks at how drilling time and production quantities differ by winner's identity in the presence of nearby production. Again, the point estimates are not statistically significant. Point estimates suggest that leases won by firms produce relatively larger quantities. This would be consistent with information asymmetry decreasing the frequency of marginal wells, but these results are difficult to interpret in the context of the coefficients from the first row. Taken together, we do not believe we have enough precision for this set of results to be substantively interpreted.

[1]One explanation could be crowding out. However, leases won by firms do not appear to have drilled wells more frequently.

Table A.4: Regression Results: Drilling Times & Production Quantities

	(1) Drill Time (Log Days)	(2) Drill Time (Log Days)	(3) Oil Production First 36 Mths. (Log Barrels)	(4) Oil Production First 36 Mths. (Log Barrels)
Firm Winner	-0.287	-0.060	-1.345	-1.080
	(0.251)	(0.231)	(1.391)	(0.834)
Nearby Production Flag	-0.207	-0.006	-0.244	0.274
	(0.159)	(0.088)	(0.810)	(0.306)
Firm/Nearby Prod Interaction	0.051	0.009	1.174	0.405
	(0.323)	(0.355)	(1.404)	(1.064)
Full Sample	No	Yes	No	Yes
Offers & Acreage Controls	Yes	Yes	Yes	Yes
Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.089	0.113	0.195	0.237
Observations	156	1241	91	635

Notes: Columns 1 and 2 look at the length of time it takes to drill an oil well. Durations longer than 180 days or less than 1 day are excluded because they indicate a data error or that there was likely a stoppage in the middle of drilling. Columns 3 and 4 look at oil production on productive wells. Columns 1 and 3 use our restricted sample with exactly one firm winner. Columns 2 and 4 use our full sample, relying on our controls to correct for endogenous entry. All specifications look at wells drilled within twelve years of the lease date.

A.6 Alternative Nearby Producing Firm Trades Analysis

This appendix includes an alternative set of nearby trades. In contrast to our preferred specification, we only include trades that happened before any wells are drilled on a lease and before twelve years have elapsed. Results are similar to those in our preferred specification, Table 1.6.

Table A.5: Regression Results: NPF Trades, Alternative Definition

	Sample with Nearby Firms		Sample with Distant Firms	
	(1)	(2)	(3)	(4)
Firm Winner	-0.054**	-0.052***	-0.023	0.001
	(0.026)	(0.016)	(0.023)	(0.019)
intercept	0.047*	0.030	0.041*	-0.002
	(0.025)	(0.020)	(0.023)	(0.013)
Full Sample	No	Yes	No	Yes
Offers & Acreage Controls	Yes	Yes	Yes	Yes
Lottery Year Fixed Effects	No	Yes	No	Yes
R squared	0.024	0.019	0.038	0.013
Observations	376	2399	443	2644

Notes: We examine how frequently leases are reassigned to firms that have nearby production. Column (1) uses our restricted sample and Column (2) uses the full sample. Columns (3) and (4) look at how frequently leases are reassigned to distant firms with production in the area. A pooled regression testing whether the 'Firm Winner' coefficient is different in columns (1) and (3) yields a p-value of 0.363. A pooled regression testing whether the 'Firm Winner' coefficient is different in columns (2) and (4) yields a p-value of 0.038. This differs from our preferred specification by requiring the matched reassignment to happen before drilling began on a lease and before twelve years.

APPENDIX B

Second Chapter Appendices

B.1 Variation in Gas Price Data

Much of the short-term variation in the price of natural gas is driven by demand shocks due to changes in temperature or forecasted temperature. Home heating is the largest natural gas demand source in the United States. During especially cold winters, gas use and gas prices rise. The reverse is true during warmer winters. To understand this relationship, I use seven-day average future continental US heating degree days (HDD) provided by the National Weather Service. I use continental (as opposed to interconnection-level) HDD because the natural gas market is national – very cold temperatures in the Northeast will cause gas prices in Texas to increase.

This analysis does not capture the full effect of forecasted temperature on natural gas prices. However, it does show that the relationship is strong. Future heating degree days will not capture gas price variation caused by unusually warm temperatures (which cause electricity demand to increase, increasing gas demand). Additionally, I am using actual future HDD as a proxy for forecasted HDD. Note that this analysis does not include variation caused by short-run supply changes.

I now conduct a Frisch-Waugh decomposition to better understand the variation in the gas price data. Specifically, I regress the residuals from each of the following equations on each other:

$$\begin{aligned} P_t^{NG} = \alpha_0 + s(Q_t^E) + \mathbb{1}\{P_t^{NG} > med(P_t^{NG})\} * s(Q_t^E) + s(Renewables_t) \\ + s(HDD_t^{current}) + s(CDD_t^{current}) + s(Date_t) + \gamma D_m + \epsilon_t \end{aligned} \tag{B.1}$$

$$HDD_t^{future} = \alpha_0 + s(Q_t^E) + \mathbb{1}\{P_t^{NG} > med(P_t^{NG})\} * s(Q_t^E) + s(Renewables_t) + s(HDD_t^{current}) + s(CDD_t^{current}) + s(Date_t) + \gamma D_m + \epsilon_t \quad \text{(B.2)}$$

The dependent variables here are the same as those in equation 2.2. Once I control for seasonal variation, time trends, renewable generation, and electricity demand, I am able to look at how residual variation in weather forecasts affects variation in the price of natural gas. My standard errors are calculated as in other parts of the paper; I use a Newey-West specification with seven lags.

Figure B.1 displays this relationship for selected hours. Even though the average future continental HDD is a crude measure of the effect of weather on gas prices, the relationship is still strong. Forecasts with above expectation HDD (that is, unusually cold days) are associated with gas prices that are above expectation. Similarly, unusually warm days are associated with gas prices that are below expectation. This is the relationship that I expect. All 72 interconnection-hour combinations are significant at the 10% level; the t-statistics range between 1.78 and 3.43. Six t-statistics fall between 1.78 and the 5% threshold of 1.96; all six are in the Texas interconnection.

B.2 Alternative Construction Regressions

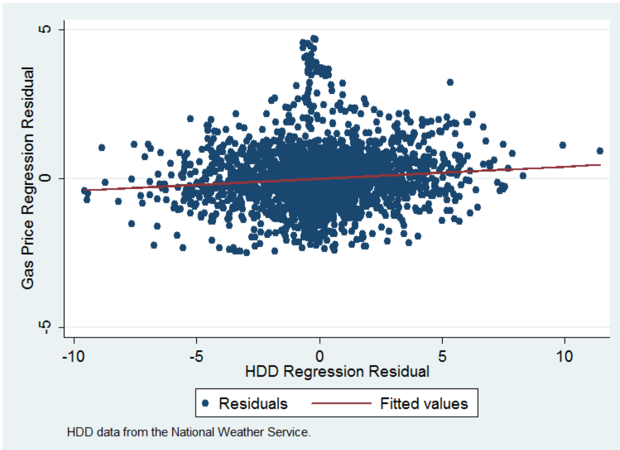
This section expands on the construction regressions in section 2.3.1. I first present the data in scatterplot form and include a line of best fit (Figure B.2). The relationship appears negatively correlated for both sources of estimated construction starts data.

I also present alternative specifications that have leads of between 12 and 36 months. I use a minimum 12-month lead, which is shorter than an 18-month construction period, to capture the fact that some paused construction could have been restarted with the advent of cheaper natural gas. The EIA-860 micro data supports this possibility; many projects are listed as "Indefinitely Postponed."[1] It is highly likely that some of these postponed projects were continued after gas prices dropped. I use a maximum 36-month lead because it is the longest planned construction duration for a gas-fired power plant. Additionally, my data does not allow me to analyze longer-term construction effects.

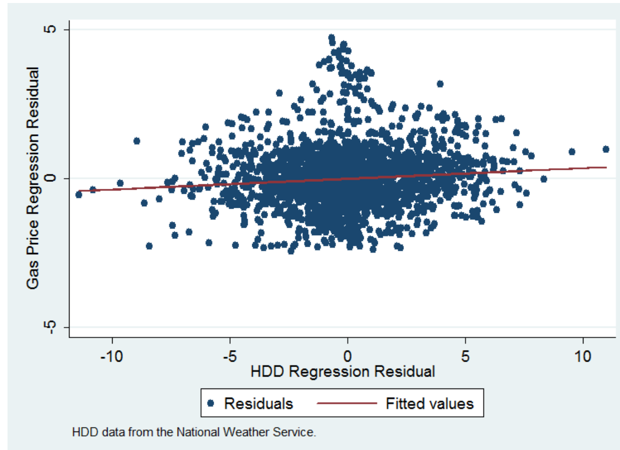
I consider both logged and raw construction starts as the independent variable. Tables B.1 and B.2 display the results of alternative specifications. Columns [2] and [7] of Table B.2 are my preferred specifications.

[1]The results using AEO data with a 12-month lead and a logged dependent variable do not strongly support this possibility.

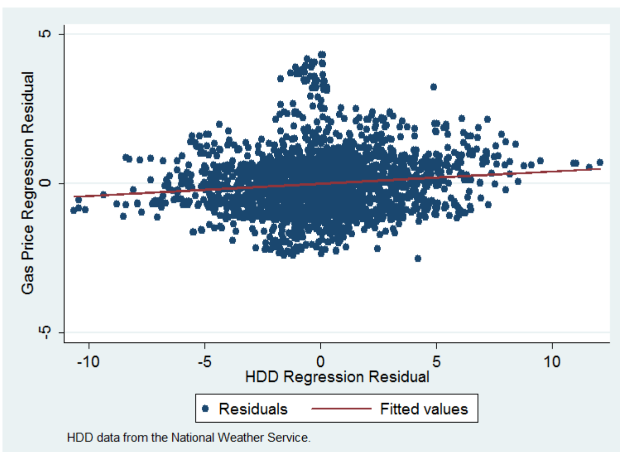
Figure B.1: Frisch-Waugh Analysis of Variation



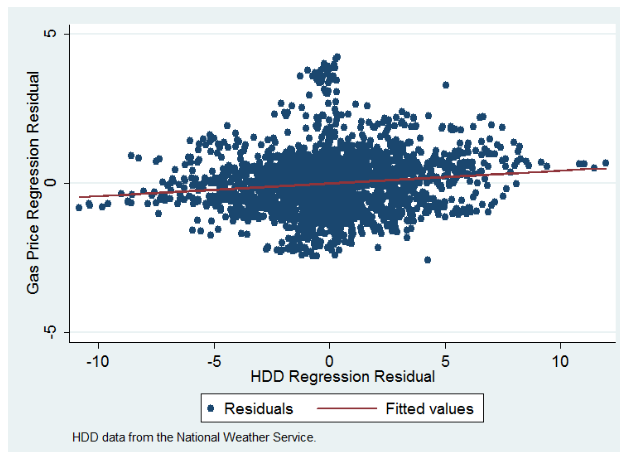
(a) Eastern: 2:00 AM (Off-Peak)



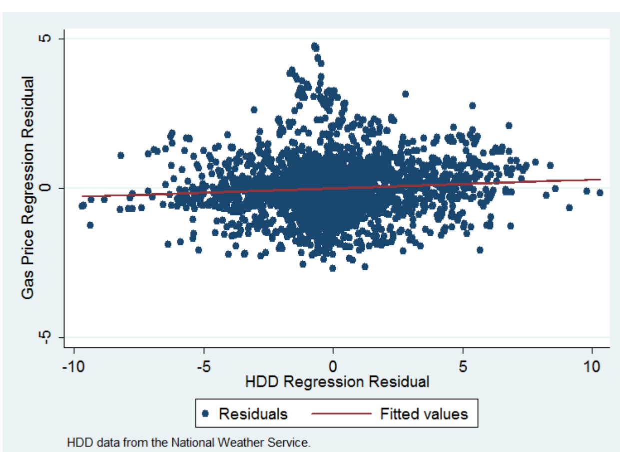
(b) Eastern: 6:00 PM (Peak)



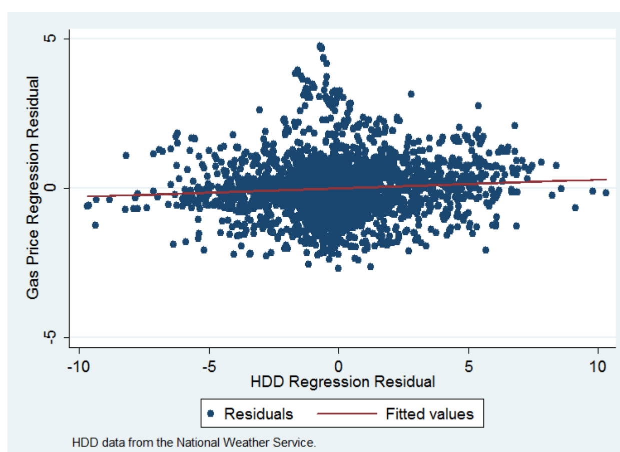
(c) Western: 2:00 AM (Off-Peak)



(d) Western: 6:00 PM (Peak)



(e) Texas: 2:00 AM (Off-Peak)



(f) Texas: 6:00 PM (Peak)

Table B.1: Regression Results: Counterfactual Plant Construction, Raw Starts

Estimated Counterfactual Plant Construction
Results from Regression Analysis Using Raw Starts

Item		Using AEO Plant Data					Using EIA-860 Plant Data				
		12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag	12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag
		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Gas Price Coefficient	[a]	-1.34	-2.57	-2.79	-2.90	-2.09	-3.47	-4.11	-3.69	-4.48	-3.24
Standard Error	[b]	2.57	2.35	2.79	2.37	1.89	2.68	2.65	2.86	2.39	2.03

Counterfactual Plant Construction (Gigawatts)

Year		12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag	12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag
2009	[c]	10.9	10.9	10.9	10.9	10.9	9.4	9.4	9.4	9.4	9.4
2010	[d]	-4.4	-4.1	2.3	2.3	2.3	-10.7	-3.7	6.5	6.5	6.5
2011	[e]	3.3	-2.8	-4.4	2.3	9.5	-6.3	-10.0	-8.7	-1.5	9.7
2012	[f]	0.7	-4.9	-5.3	-6.3	-2.8	-8.5	-10.7	-7.8	-12.3	-6.9
2013	[g]	-2.0	-8.1	-7.6	-7.5	-3.0					
Actual Construction 2010 - 2012/2013	[h]	25.9	25.9	25.9	25.9	25.9	25.4	25.4	25.4	25.4	25.4
Construction Caused by Fracking	[i]	**21.9**	**25.9**	**23.6**	**21.3**	**14.1**	**25.4**	**25.4**	**18.9**	**18.9**	**9.2**

Notes: EIA-860 data for 2013 is unavailable. Gas Plants take between 18 and 36 months to construct.
[i]: = [h] - Sum([d] to [g]) if positive & shaded.
Sources: EIA-860 data summarized in the Electric Power Annual and Annual Energy Outlook Data.

Table B.2: Regression Results: Counterfactual Plant Construction, Logged Starts

Estimated Counterfactual Plant Construction
Results from Regression Analysis Using Logged Construction Starts

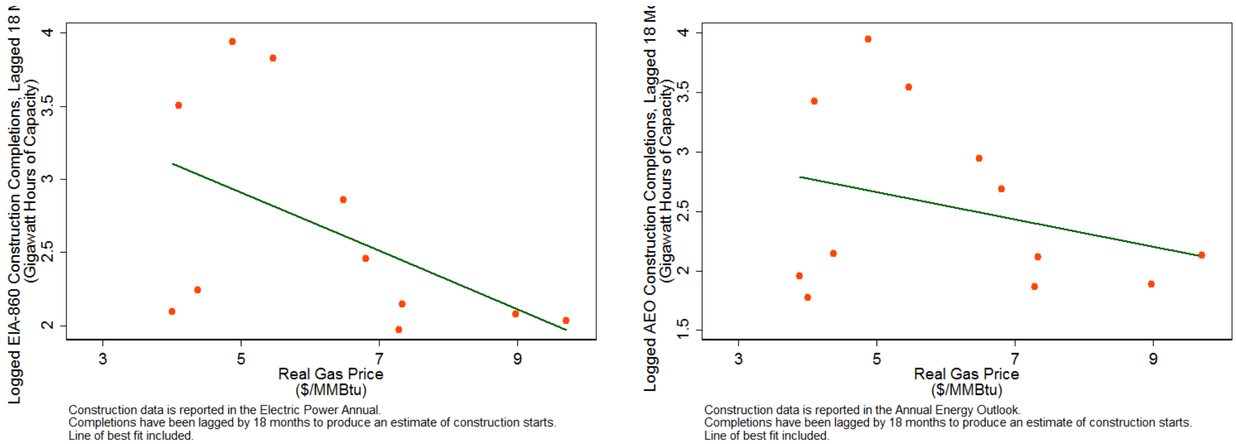
Item		Using AEO Plant Data					Using EIA-860 Plant Data				
		12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag	12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag
		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Gas Price Coefficient	[a]	-0.02	-0.13	-0.21	-0.19	-0.16	-0.15	-0.21	-0.21	-0.25	-0.18
Standard Error	[b]	0.13	0.12	0.13	0.11	0.13	0.12	0.12	0.11	0.10	0.10

Counterfactual Plant Construction (Gigawatts)

Year		12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag	12-Month Lag	18-Month Lag	24-Month Lag	30-Month Lag	36-Month Lag
2009	[c]	10.9	10.9	10.9	10.9	10.9	9.4	9.4	9.4	9.4	9.4
2010	[d]	2.0	1.5	2.3	2.3	2.3	1.7	3.2	6.5	6.5	6.5
2011	[e]	8.5	3.6	-0.4	5.1	9.5	3.0	0.1	-0.5	3.7	9.7
2012	[f]	6.7	2.8	0.3	0.8	1.7	2.2	0.0	0.2	-1.8	0.8
2013	[g]	5.7	1.7	-0.5	0.6	1.8					
Actual Construction 2010 - 2012/2013	[h]	25.9	25.9	25.9	25.9	25.9	25.4	25.4	25.4	25.4	25.4
Construction Caused by Fracking	[i]	**2.9**	**16.2**	**23.4**	**17.2**	**10.7**	**18.5**	**22.1**	**18.7**	**15.2**	**8.5**

Notes: EIA-860 data for 2013 is unavailable. Gas Plants take between 18 and 36 months to construct.
[i]: = [h] - Sum([d] to [g]) if positive & shaded.
Sources: EIA-860 data summarized in the Electric Power Annual and Annual Energy Outlook Data.

Figure B.2: Gas Prices and Estimated Construction Starts



(a) Estimated EIA-860 Construction Starts

(b) Estimated AEO Construction Starts

B.3 Primary Specification Fit

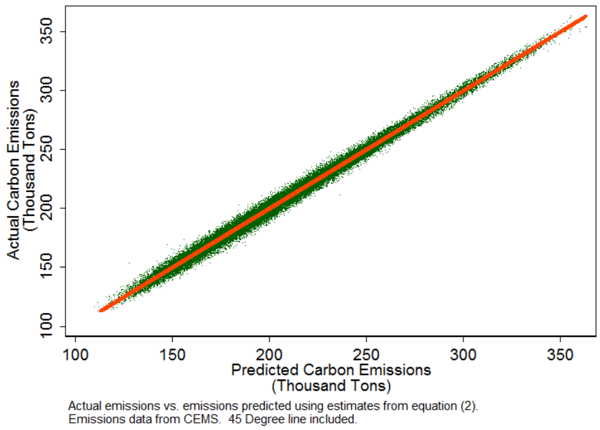
Figure B.3 displays scatter plots for each interconnection with a $45°$ line. It maps actual carbon emissions against carbon emissions predicted by my primary specification. The relationship is very strong and most points are very close to the $45°$ line. There does not appear to be any bias at the aggregate level.

B.4 Alternative Gas Prices

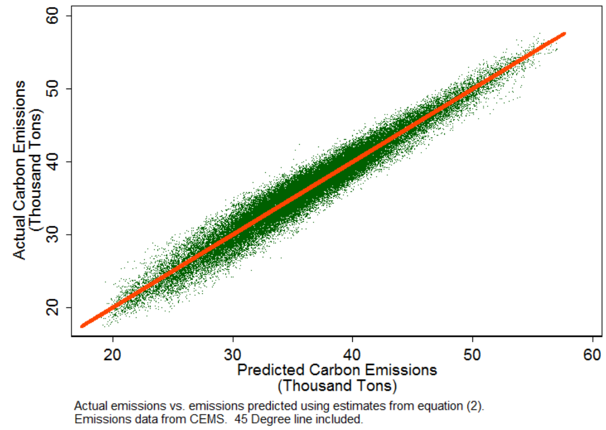
This paper uses natural gas prices at Henry Hub, the major US trading location in southern Louisiana. Using Henry Hub gas prices has several advantages. Henry Hub has been a consistent trading location for decades and most other locations trade with prices marked to Henry Hub. In contrast, many other hubs have thin trading at some times and/or can change exact terms and locations of delivery over time. Additionally, it provides a benchmark that allows for easy interpretation of the results and extrapolation. While Eastern and Texas gas prices are very closely linked to Henry Hub, Western gas prices occasionally are out of sync. This subsection analyzes the Western interconnection using gas prices at SoCal Gas, a major trading hub in California.

Figure B.4 plots gas prices at Henry Hub (in the Eastern interconnection), Katy (in Texas), and SoCal (in the WECC). Generally, the three prices closely track each other, though SoCal is a looser fit. In particular, late 2008 saw SoCal deviate from the Henry Hub/Katy price. The daily correlation between Henry Hub and Katy gas prices is .996, while the daily correlation between

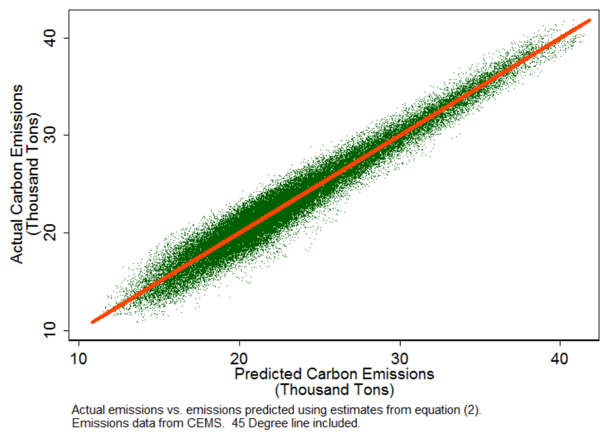
Figure B.3: Actual Emissions vs. Predicted Emissions



(a) Eastern Interconnection



(b) Western Interconnection



(c) Texas Interconnection

Henry Hub and SoCal is .973.

I rerun my analysis with the only change being the different gas price for the Western interconnection. Figure B.5 highlights two of the key results. The gas price spline in panel (a) looks similar to the spline using Henry Hub prices in Figure 2.7. Counterfactual emissions estimates in panel (b) look similar to counterfactual estimates using Henry Hub prices in Figure 2.8. Total 2013 reductions using prices at SoCal for the Western interconnection are estimated at 15.5 thousand tons/hour. Using Henry Hub gas prices for all interconnections, total 2013 reductions are estimated at 16.7 million tons/hour. The difference is primarily driven by the fact that late 2008 gas prices at SoCal were lower than at Henry Hub. Additionally, the end of 2013 saw higher prices at SoCal than at Henry Hub.

Figure B.4: Seven-Day Moving Average Gas Prices

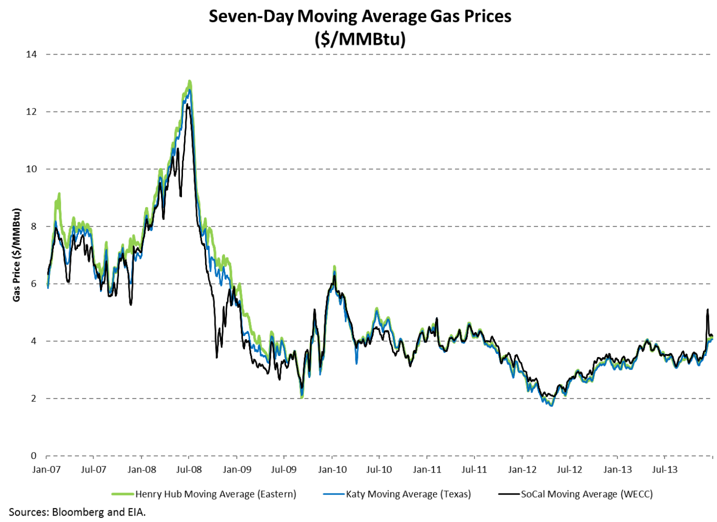
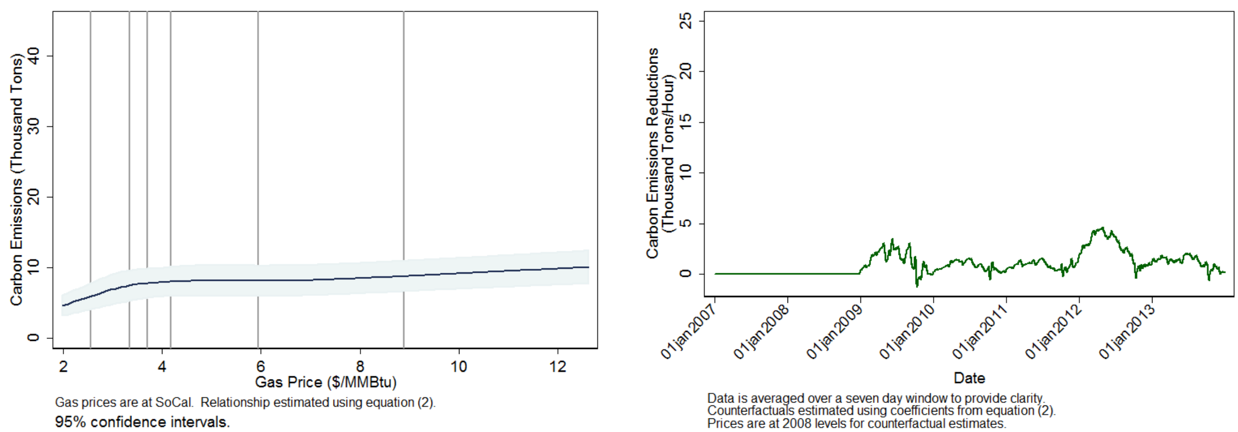


Figure B.5: Western Interconnection using Gas Prices at SoCal Trading Hub



(a) WECC 6:00 PM (Peak)

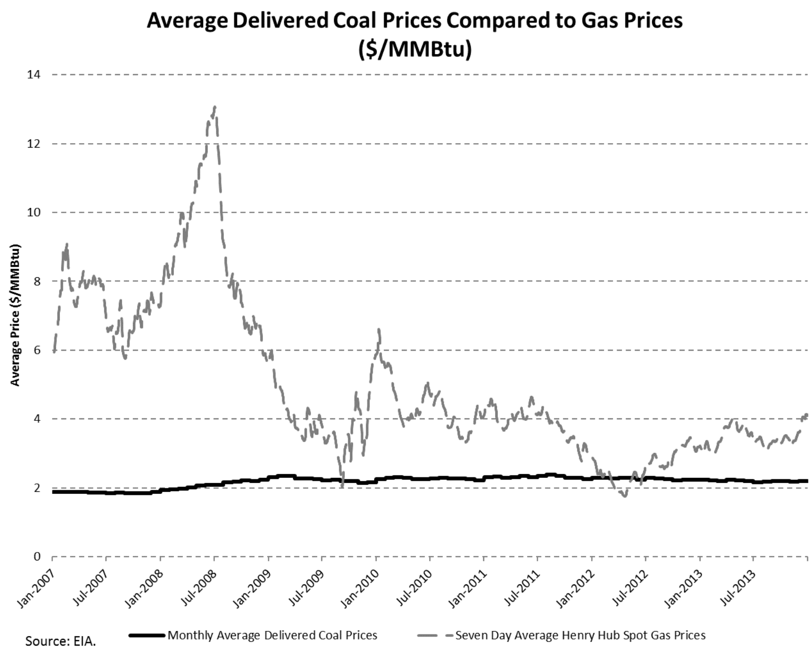
(b) WECC Counterfactuals using SoCal Prices

B.5 Inclusion of Coal Prices

Coal prices are an important component of the electricity market. While my preferred specification omits them because natural gas prices influence coal prices, I consider them here. I use data from the EIA's Short-Term Energy Outlook. The EIA constructs an average price for coal delivered to electricity plants across the country. I use the delivered price because trading hubs are generally not close to coal-fired generators.

Coal is very expensive to transport, so the spot price in, e.g., Central Appalachia does a poor job of approximating the marginal cost of burning coal for most generators (Cicala, 2014). Delivered coal prices are only available on a monthly basis. This should not be a problem because most coal contracts are medium-term and coal generators are not easily able to resell coal on the spot market because of the transportation costs.[2] Figure B.6 plots the monthly average delivered coal price from 2007 to 2013.[3] I include the seven-day average Henry Hub spot price of natural gas for comparison. Delivered coal prices are much more stable than natural gas prices.

Figure B.6: Average Delivered Coal Prices Compared to Gas Prices



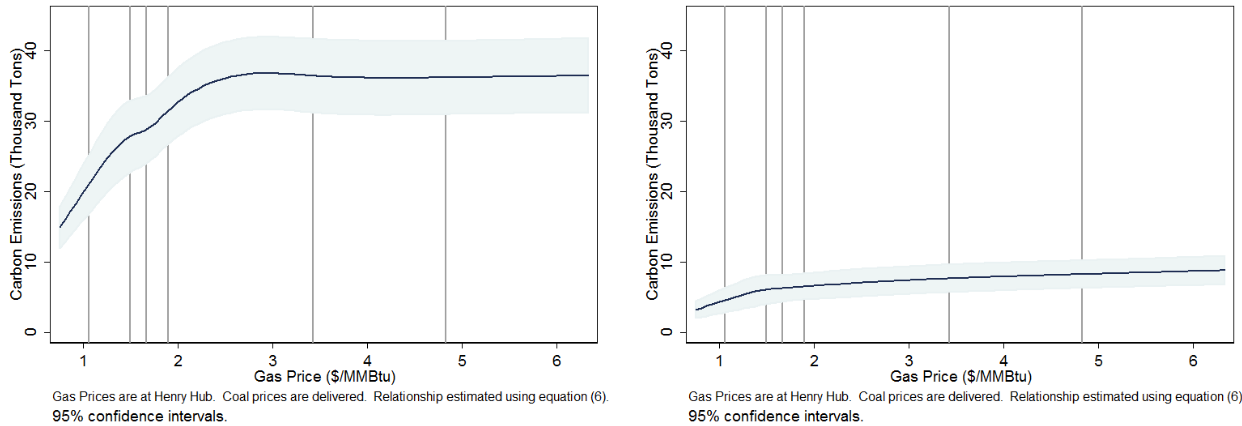
Specifically, I estimate:

$$TE_t = \alpha_0 + \mathbf{s}(\mathbf{P_t^{NG}}/\mathbf{P_t^{Coal}}) + s(Q_t^E) + \mathbb{1}\{P_t^{NG} > med(P_t^{NG})\} * s(Q_t^E) + s(Renewables_t)$$
$$+ s(HDD_t) + s(CDD_t) + s(Date_t) + \gamma D_m + \epsilon_t \quad \text{(B.3)}$$

When estimating counterfactual emissions, I estimate all emissions as if both 2008 gas prices and 2008 coal prices were realized in every other year. Figure B.7 shows the resulting gas/coal price ratio splines for the Eastern and Western interconnections at 6:00 PM. I omit the Texas interconnection (as well as other parts of analysis) for brevity; omitted results look similar to my primary specification. Delivered coal prices were fairly constant at a little more than $2/MMBtu dur-

[2]Coal plants adjust production by storing inventory at low cost and purchasing less coal in future contracts.
[3]Delivered coal prices are reported at the monthly level.

Figure B.7: Results using a Gas/Coal Price Ratio



(a) Eastern Price Ratio Spline: 6:00 PM

(b) Western Price Ratio Spline: 6:00 PM

ing this time period, indicating that a price ratio of 1 usually has gas prices just above $2/MMBtu. The splines look very similar to corresponding splines from my preferred specification in Figure 2.7.

Table B.3 displays counterfactual estimates. Results look similar to counterfactual estimates using only gas prices in Table 2.4. Total estimated reductions in 2013 using my primary analysis are 16.7 thousand tons of carbon per hour. When I use a gas/coal price ratio, counterfactual emissions are estimated at 16.3 thousand tons of carbon per hour (s.e. of 1.0). 2013 reductions attributable to new plant construction are 2.1 thousand tons/hour in my primary specification, and remain 2.1 thousand tons/hour when coal prices are included (s.e. of 0.06).

Table B.3: Hourly Reductions in Carbon Dioxide Emissions, Gas/Coal Price Ratio

Hourly Reductions in Carbon Dioxide Emissions
By Year - Using Gas/Coal Price Ratio
(Thousand Tons of Carbon Dioxide/Hour)

		2008	2009	2010	2011	2012	2013
		[1]	[2]	[3]	[4]	[5]	[6]
Actual Emissions	[a]	284.7	261.3	278.3	266.0	249.0	253.0
Predicted Emissions using Gas/Coal Price Ratio	[b]	284.3	261.9	277.6	266.4	249.1	252.8
Counterfactual Emissions (Using 2008 Gas Prices)	[c]		274.0	287.6	279.8	271.6	267.2
Emissions Reductions From Gas/Coal Switching	[d]		12.1	10.0	13.4	22.5	14.4
			(0.9)	(0.9)	(1.0)	(1.2)	(1.0)
Counterfactual Emissions (Without Newly Constructed Plants)	[e]			277.6	267.0	250.9	254.9
Reduction Caused by New Power Plants	[f]			0.1	0.6	1.8	2.1
				(0.0)	(0.0)	(0.1)	(0.1)
Counterfactual Emissions (With 2008 Gas Prices and Without New Plants)	[g]		274.0	287.7	280.3	273.3	269.1
Total Reductions from Low Gas Prices	[h]		12.1	10.1	13.9	24.2	16.3
			(0.9)	(0.9)	(1.0)	(1.2)	(1.0)

Notes:
[d] = [c] - [b]
[f] = [e] - [b]
[h] = [g] - [b]
Standard errors are estimated using block bootstrapping with 1000 replications.

APPENDIX C

Third Chapter Appendices

C.1 Additional Regression Results

This appendix presents alternative regression results. Table C.1 presents results that are non-linear in the number of competitors. Table C.2 displays results using jet fuel prices that have been lagged by two months. Results are very similar to our primary results. Table C.3 logs the dependent variable and fuel cost variables.

Table C.4 presents separate results during the carbon tax period and outside the carbon tax period. There is anecdotal evidence that the imposition of the carbon tax prompted a change in the competitive nature of the Australian airline industry.[1] Regression results support this hypothesis.

Table C.5 is similar to Table 3.4, but it includes a term that interacts the cost of jet fuel and the average capacity factor. Tables C.6 and C.7 mimic Tables 3.8 and 3.9, respectively, but they include a term that interacts the cost of jet fuel and the average capacity factor.

[1] For example, see articles in The Sydney Morning Herald and the Australian Broadcast Corporation.

Table C.1: Regression Results: Non-Linear Specification

Weighted by Number of Passengers

	Base	O-D FE	More F-E	State/Route Controls
Number of Competitors	-4.268*			
	(2.320)			
Jet Fuel Cost	1.102***	1.252***	1.158***	1.044***
	(0.190)	(0.203)	(0.183)	(0.192)
Jet Fuel * Competitors	0.251***			
	(0.072)			
1 Competitors		-23.358***	-30.422***	-30.045***
		(3.254)	(2.879)	(3.068)
2 Competitors		-23.374***	-28.376***	-29.338***
		(4.685)	(4.045)	(3.514)
3 Competitors		-75.798***	-76.816***	-79.267***
		(6.447)	(6.424)	(6.927)
4 Competitors		-58.215***	-53.043***	-64.681***
		(7.365)	(7.256)	(7.830)
5 Competitors		9.740	-12.629	0.131
		(22.593)	(23.794)	(22.322)
1 Comps * Fuel Cost		0.002	0.237***	0.214***
		(0.079)	(0.075)	(0.076)
2 Comps * Fuel Cost		0.407**	0.522***	0.583***
		(0.168)	(0.147)	(0.139)
3 Comps * Fuel Cost		1.588***	1.540***	1.544***
		(0.165)	(0.170)	(0.171)
4 Comps * Fuel Cost		1.571***	1.287***	1.327***
		(0.340)	(0.338)	(0.328)
5 Comps * Fuel Cost		-0.709	-2.950**	-3.699***
		(1.259)	(1.382)	(1.322)
Capacity Factor	57.264***	36.969***	35.616***	23.528**
	(8.726)	(6.886)	(9.129)	(9.074)
Thousand KM, Avg	83.478***	75.226***	62.126***	77.811***
	(11.089)	(3.712)	(3.210)	(3.720)
Distance Squared	-12.834***	-1.651***	-1.649***	-2.972***
	(2.108)	(0.588)	(0.503)	(0.464)
Quarter-Year Fixed Effects	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	Yes	Yes	Yes	Yes
Origin Fixed Effects	Yes	Yes	Yes	Yes
Destination Fixed Effects	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	Yes	No	Yes	Yes
Airline Fixed Effects	Yes	No	Yes	Yes
State/Route Controls	Yes	No	No	Yes
R squared	0.833	0.771	0.787	0.792
Observations	102168	112965	112965	102297

Notes: The dependent variable is average airfare at the month-airline-origin-destination level. For example, one observation is all passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation.

Table C.2: Regression Results: Jet Fuel Prices Lagged by Two Months

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls
Number of Competitors	-42.554***	-21.788***	-7.008***	-7.634***	-9.361***
	(2.140)	(1.573)	(1.875)	(1.708)	(1.658)
Jet Fuel Cost	2.008***	2.367***	1.100***	1.211***	1.079***
	(0.108)	(0.134)	(0.168)	(0.155)	(0.161)
Jet Fuel * Competitors	1.969***	0.722***	0.295***	0.290***	0.346***
	(0.107)	(0.076)	(0.060)	(0.054)	(0.053)
Capacity Factor		-97.134***	43.326***	41.460***	27.749***
		(7.356)	(6.619)	(8.720)	(8.692)
Thousand KM, Avg		50.658***	70.964***	58.046***	73.907***
		(2.941)	(3.408)	(3.062)	(3.497)
Distance Squared		-6.475***	-1.393**	-1.369**	-2.633***
		(0.298)	(0.571)	(0.520)	(0.495)
Quarter-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes
State/Route Controls	No	No	No	No	Yes
R squared	0.501	0.632	0.768	0.785	0.790
Observations	112965	112965	112965	112965	102297

Notes: The dependent variable is average airfare at the month-airline-origin-destination level. For example, one observation is all passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by two months. Weights are assigned according to the number of passengers in an observation.

Table C.3: Regression Results: Log-Log Specification

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls
Number of Competitors	-0.349***	-0.280***	-0.253***	-0.139***	-0.151***
	(0.028)	(0.021)	(0.022)	(0.020)	(0.022)
Log Jet Fuel Cost	0.351***	0.241***	0.239***	0.091***	0.068***
	(0.008)	(0.015)	(0.011)	(0.013)	(0.015)
Log Jet Fuel * Competitors	0.108***	0.083***	0.081***	0.042***	0.045***
	(0.009)	(0.006)	(0.007)	(0.006)	(0.006)
Capacity Factor		-0.553***	-0.610***	0.124***	0.094*
		(0.050)	(0.062)	(0.046)	(0.048)
Thousand KM, Avg		0.149***	0.145***	0.321***	0.393***
		(0.013)	(0.010)	(0.016)	(0.019)
Distance Squared		-0.011***	-0.011***	-0.031***	-0.038***
		(0.001)	(0.001)	(0.002)	(0.003)
Quarter-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	No	Yes	Yes
Destination Fixed Effects	No	No	No	Yes	Yes
Num Stops Fixed Effects	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes
R squared	0.447	0.570	0.609	0.725	0.728
Observations	112482	112482	112482	112482	102132

Notes: The dependent variable is the log of average airfare at the month-airline-origin-destination level. For example, one observation is all passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation.

Table C.4: Regression Results: Inside/Outside of Carbon-Tax Period

Weighted by Number of Passengers

	Base	Outside Tax	Tax Period
Number of Competitors	-9.302***	-5.614***	-22.422***
	(1.696)	(1.756)	(2.440)
Jet Fuel Cost	1.011***	1.838***	-0.568**
	(0.173)	(0.283)	(0.267)
Jet Fuel * Competitors	0.339***	0.263***	0.457***
	(0.053)	(0.055)	(0.087)
Capacity Factor	28.067***	25.132**	26.873**
	(8.716)	(12.186)	(9.971)
Thousand KM, Avg	75.144***	52.809***	98.615***
	(3.578)	(4.545)	(4.437)
Distance Squared	-2.617***	0.370	-2.991***
	(0.493)	(0.477)	(0.529)
Quarter-Year Fixed Effects	Yes	Yes	Yes
Month-of-Year Fixed Effects	Yes	Yes	Yes
Origin Fixed Effects	Yes	Yes	Yes
Destination Fixed Effects	Yes	Yes	Yes
Num Stops Fixed Effects	Yes	Yes	Yes
Airline Fixed Effects	Yes	Yes	Yes
State/Route Controls	Yes	Yes	Yes
R squared	0.790	0.779	0.856
Observations	102297	65201	37096

Notes: The dependent variable is the average airfare at the month-airline-origin-destination level. For example, one observation is all passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Regression run separately for tax- and no-tax- periods.

Table C.5: Regression Results: Observations separated by Number of Stops, Including Capacity Interaction

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls
Number of Competitors	-40.651***	-21.567***	-7.050***	-6.575***	-8.136***
	(1.976)	(1.505)	(1.906)	(1.853)	(1.719)
Jet Fuel Cost	0.026	1.391***	0.501**	0.205	-0.210
	(0.336)	(0.206)	(0.222)	(0.247)	(0.255)
Jet Fuel * Competitors	1.855***	0.726***	0.263***	0.221***	0.276***
	(0.097)	(0.078)	(0.061)	(0.057)	(0.054)
Capacity Factor	-123.514***	-110.290***	29.345***	18.487**	-2.966
	(11.166)	(6.903)	(5.814)	(7.187)	(7.137)
Jet Fuel * Capacity	2.931***	0.935***	0.619**	1.080***	1.488***
	(0.533)	(0.274)	(0.297)	(0.355)	(0.371)
Thousand KM, Avg		51.758***	82.324***	69.571***	81.596***
		(2.716)	(2.780)	(2.619)	(3.218)
Distance Squared		-5.456***	-3.619***	-3.137***	-3.407***
		(0.258)	(0.182)	(0.183)	(0.203)
Quarter-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes
State/Route Controls	No	No	No	No	Yes
R squared	0.500	0.622	0.759	0.772	0.779
Observations	169394	169394	169394	169394	156573

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-number of stops level. For example, one observation is all non-stop passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation.

Table C.6: Regression Results: Economy Class, Including Capacity Interaction

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	-41.474***	-23.725***	-6.588***	-7.638***	-9.631***	-4.587**
	(2.066)	(1.644)	(1.931)	(1.770)	(1.720)	(2.202)
Jet Fuel Cost	1.987***	2.140***	0.640**	0.374	-0.098	0.274
	(0.108)	(0.440)	(0.263)	(0.289)	(0.303)	(0.293)
Jet Fuel * Competitors	1.954***	0.957***	0.303***	0.299***	0.358***	0.262***
	(0.112)	(0.102)	(0.067)	(0.062)	(0.062)	(0.078)
Capacity Factor		-119.154***	30.708***	20.316***	-2.492	38.522***
		(10.311)	(5.883)	(7.456)	(7.747)	(6.761)
Jet Fuel * Capacity		1.603***	0.560*	1.040**	1.520***	1.004**
		(0.408)	(0.320)	(0.395)	(0.417)	(0.382)
Thousand KM, Avg		29.439***	69.562***	57.082***	72.855***	71.140***
		(4.163)	(3.327)	(3.056)	(3.531)	(10.721)
Distance Squared		-6.218***	-1.691***	-1.551***	-2.864***	-10.743***
		(0.363)	(0.520)	(0.480)	(0.458)	(1.996)
Quarter-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.466	0.513	0.752	0.768	0.773	0.821
Observations	118779	118779	118779	118779	107798	107665

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-cabin class level. For example, one observation is all economy class passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Only economy class fares are included.

Table C.7: Regression Results: Business Class, Including Capacity Interaction

Weighted by Number of Passengers

	No Fixed Effects	Mileage	O-D FE	More FE	State/Route Controls	Route FE
Number of Competitors	51.242***	49.206***	20.959**	21.053**	23.195**	9.491
	(9.801)	(11.195)	(9.898)	(9.299)	(9.220)	(9.025)
Jet Fuel Cost	5.154***	2.495***	2.348***	2.398***	2.396***	2.002**
	(0.276)	(0.871)	(0.788)	(0.801)	(0.801)	(0.842)
Jet Fuel * Competitors	-1.001***	-0.958***	-0.436**	-0.286	-0.152	0.132
	(0.277)	(0.288)	(0.189)	(0.181)	(0.186)	(0.190)
Capacity Factor		-136.674**	221.727***	214.724***	69.937	106.288
		(52.456)	(46.888)	(59.761)	(62.031)	(64.084)
Jet Fuel * Capacity		1.370	-1.131	-0.655	-0.804	-1.075
		(1.118)	(1.043)	(1.098)	(1.102)	(1.175)
Thousand KM, Avg		110.942***	130.316***	99.738***	81.982***	749.425***
		(13.897)	(11.408)	(11.808)	(13.352)	(134.729)
Distance Squared		-6.258**	3.436*	1.375	9.322***	-110.621***
		(2.647)	(1.888)	(2.009)	(2.157)	(17.635)
Quarter-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Month-of-Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Origin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Destination Fixed Effects	No	No	Yes	Yes	Yes	Yes
Num Stops Fixed Effects	No	No	No	Yes	Yes	Yes
Airline Fixed Effects	No	No	No	Yes	Yes	Yes
State/Route Controls	No	No	No	No	Yes	Yes
R squared	0.607	0.628	0.743	0.772	0.792	0.830
Observations	22725	22725	22725	22725	21405	21301

Notes: The dependent variable is the average airfare at the month-airline-origin-destination-cabin class level. For example, one observation is all business class passengers traveling from Sydney to Melbourne on Qantas in January 2010. Competitors are at the operating-airline level, and at least 1% market share is required to be designated a competitor. Prices are in 2010 USD. Standard errors are clustered at the month level. Jet fuel prices are lagged by one month. Weights are assigned according to the number of passengers in an observation. Only business class fares are included.