Essays in Energy and Environmental Economics

by

Brian C. Prest

University Program in Environmental Policy
Duke University

Date: ____________________

Approved:

________________________
William A. Pizer, Supervisor

________________________
Richard G. Newell

________________________
Steven E. Sexton

________________________
Christopher D. Timmins

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the University Program in Environmental Policy
in the Graduate School of Duke University
2018
Abstract

Essays in Energy and Environmental Economics

by

Brian C. Prest

University Program in Environmental Policy
Duke University

Date: __________________________

Approved:

______________________________
William A. Pizer, Supervisor

______________________________
Richard G. Newell

______________________________
Steven E. Sexton

______________________________
Christopher D. Timmins

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University Program in Environmental Policy in the Graduate School of Duke University 2018
Abstract

This dissertation includes three papers discussing different aspects of environmental and energy economics and policy. The first paper (“Peaking Interest: How awareness drives the effectiveness of time-of-use electricity pricing”) analyzes what factors drive consumer responses to time-of-use electricity pricing. Such pricing is an increasingly common policy that is designed to reduce reliance on high-cost, inefficient, and often polluting power plants. Using both survey and meter data on Irish households, I find that behavioral factors drive consumer responses: consumer awareness and information are key to inducing responses, and marginal prices have minimal effects.

The second paper (“Prices versus Quantities with Policy Updating”, with William Pizer) is a theoretical study considering how intertemporal trading (“banking”) under cap-and-trade style policies changes the traditional trade-off between price and quantity regulation (Weitzman 1974). Our theoretical model illustrates an unappreciated advantage of quantity regulation: expected policy updates are incorporated immediately through secondary markets via a no arbitrage condition. This helps us understand the welfare implications of observed price volatility in permit markets and has implications for the design of environmental regulations, such as the debate over the relative merits of a carbon tax versus “cap and trade” policy.

The third paper (“Trophy Hunting vs. Manufacturing Energy: The Price Responsiveness of Shale Gas”, with Richard Newell and Ashley Vissing) uses well-level data to estimate how the shale revolution has changed the price responsiveness of
U.S. natural gas supply. We find that U.S. natural gas supply is approximately three times more price responsive as a result of the shale revolution, owing entirely to higher production per well. This improves our understanding of the market dynamics around natural gas supply.

Each of these papers has implications for environmental and energy policy. The first paper aims to understand how to improve the effectiveness of time-of-use electricity pricing policies. The second paper addresses an important feature of climate policy design, given high observed volatility in carbon allowance prices in the United States and Europe. The third paper aims to improve understanding of the recent changes in U.S. natural gas markets, with important implications for the fuel mix in electricity generation (in particular, coal-fired versus gas-fired generation) and hence CO₂ emissions.
Dedication

to Sarah, for supporting me for all these years, and my parents, Art and Carole Prest,
for their enduring encouragement
# Contents

Abstract iv  
List of Tables xi  
List of Figures xii  
List of Abbreviations xiv  
Acknowledgments xvi  

1 Introduction 1  

2 Peaking Interest: How awareness drives the effectiveness of time-of-use electricity pricing 3  
  2.1 Introduction and Motivation ........................................ 3  
  2.2 Literature ............................................................ 10  
  2.3 Data ................................................................. 13  
    2.3.1 Description of the Program ................................... 13  
    2.3.2 Treatment/Control Balance ................................. 17  
  2.4 Average Treatment Effects Using Difference-in-Differences ........ 20  
    2.4.1 Half-Hourly Average Treatment Effects .................. 20  
  2.5 Heterogeneity in Treatment Effects ............................... 27  
    2.5.1 Method ........................................................ 27  
    2.5.2 Causal Tree Estimation Results ............................. 35  
    2.5.3 Treatment Effects by Tariff Levels ......................... 39
### 2.5.4 Who Is Likely to Be Aware? 45

2.6 Conclusion 48

### 3 Prices versus Quantities with Policy Updating
(Co-author: William A. Pizer) 52

3.1 Introduction 52

3.2 The Weitzman Result and Subsequent Work 59

3.3 Policy Updating 63

3.3.1 Firm Behavior 66

3.3.2 Optimal Price and Quantity Policies 67

3.3.3 Comparative Advantage 70

3.3.4 Discussion of Key Assumptions 71

3.4 Policy Updating with Noise 73

3.5 Multiple Periods 77

3.6 Application to Climate Change Policy 84

3.7 Conclusion 86

### 4 Trophy Hunting vs. Manufacturing Energy: The Price-Responsiveness of Shale Gas
(Co-authors: Richard G. Newell and Ashley B. Vissing) 90

4.1 Introduction 90

4.2 Literature 95

4.3 Industry Background and Data 97

4.3.1 Industry Structure 97

4.3.2 Data Sources 100

4.3.3 Data Description 103

4.4 Models and Results 108

4.4.1 Overview 108

4.4.2 Stage 1: Commence Drilling (Spud) a Well 108
4.4.3 Stage 2: Spud-to-Production Time .................................. 116
4.4.4 Stage 3: Production Profile over Time ............................... 120
4.4.5 Integrating Natural Gas Supply Stages to Measure Overall
Price Responsiveness ...................................................... 126
4.4.6 Simulation 2: Simulated Price Responses Over Time .......... 133
4.5 Conclusion .................................................................... 135

5 Conclusion .................................................................... 137

A Appendix to Chapter 2 ..................................................... 139
A.1 Experimental Details ..................................................... 139
A.2 Robustness Checks ........................................................ 139
   A.2.1 Pre-trend Check ..................................................... 139
   A.2.2 Placebo Test: Holidays and Weekends ...................... 141
   A.2.3 Time-Varying in Responses to TOU Pricing ............... 143
   A.2.4 Causal Tree with Propensity Score Weighting ............ 145
   A.2.5 Honest Tree .......................................................... 147
A.3 Confirming No Heterogeneity on Appliances and Temperature . 150
A.4 Overview of Regression Trees and Cross Validation .......... 152
A.5 Consumer Surplus ........................................................ 154
   A.5.1 Costs: Effect on Consumers .................................. 156
   A.5.2 Benefits: Capacity Savings from Reduced Peak Load ... 159
A.6 List of Survey Questions Used .......................................... 165
A.7 Summary Statistics of Selected Variables .......................... 169

B Appendix to Chapter 3 ..................................................... 170
B.1 Comparative Advantage in Two Periods ............................... 170
B.2 Multiple Periods with Noise ............................................ 171
   B.2.1 Equilibrium under a Price Policy ............................. 172
## List of Tables

2.1 Treatment and Control Group Assignments ........................................... 18
2.2 Treatment/Control Balance: Covariate t-tests ........................................ 21
2.3 Treatment/Control Balance: Linear Probability Model of Treatment on Covariates .......................................................... 22
2.4 Post-Selection Lasso Linear Probability Model of Awareness ............... 50
4.1 Summary Statistics ................................................................. 105
4.2 Drilling Estimation Results .......................................................... 115
4.3 Spud-to-Production Duration Model Results ........................................ 118
4.4 Well Production Profile Fixed Effects Regressions ............................... 123
A.1 Effect on Treated Consumers during Treatment Period (Jan-Dec 2010) 159
A.2 Estimated Benefits from Reduced Required Capacity .......................... 162
A.3 List of Survey Questions Used ...................................................... 165
A.4 Selected Summary Statistics ........................................................ 169
C.1 Revenues Model First Stage (Column 1 of Table 4.2) ......................... 187
C.2 Prices-Only First Stage (Column 2 of Table 4.2) ............................... 188
C.3 Prices Plus Productivity First Stage (Column 3 of Table 4.2) ............. 189
C.4 Unconventional-Only First Stage (Column 4 of Table 4.2) ............... 190
C.5 Conventional-Only First Stage (Column 5 of Table 4.2) .................... 191
List of Figures

2.1 Time-of-Use Pricing Structures ............................................. 17
2.2 Average Treatment Effects by Time of Day ............................... 25
2.3 Average Consumption by Time of Day, Treatment Assignment, and Period .............................................................................. 27
2.4 Heterogeneous Treatment Effects on Consumption during Peak Periods 39
2.5 Estimated Demand Curves, by Pricing Period ............................... 43
2.6 Simulated Consumption under Different Pricing Regimes: January 4, 2010 ................................................................. 45
2.7 Histograms of Fitted Probabilities of Awareness, from Models shown in Table 2.4 ................................................................. 51
3.1 History of SO$_2$ allowance spot market prices as 2010 updates were developed ................................................................. 55
3.2 Timeline of events with updating in a simple two-period model .... 65
3.3 Illustration of deadweight loss from price policy .......................... 71
3.4 Timeline of events with updating in a $T$-period model .................. 79
4.1 Illustration of the Effect of Flatter Supply on Price Volatility .......... 93
4.2 Location of Gas Wells in Data by Well Type and Selected Shale Plays 104
4.3 Distributions of Initial Gas Production (IP) by Well Type, Log Scale 105
4.4 Number of Spuds by Well Type (right axis) and Oil and Gas Prices (left axis), 2000-2015, Quarterly ................................. 107
4.5 Illustration of Gas Price Effect on Spud-to-Production Time, by Well Type ................................................................. 120
## List of Abbreviations

### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SLS</td>
<td>Two-stage least squares</td>
</tr>
<tr>
<td>AFT</td>
<td>Accelerated failure time</td>
</tr>
<tr>
<td>AR</td>
<td>Autoregressive</td>
</tr>
<tr>
<td>bcf/d</td>
<td>Billion cubic feet per day</td>
</tr>
<tr>
<td>CAIR</td>
<td>Clean Air Interstate Rule</td>
</tr>
<tr>
<td>CART</td>
<td>Classification and regression trees</td>
</tr>
<tr>
<td>CATE</td>
<td>Conditional average treatment effect</td>
</tr>
<tr>
<td>CDD</td>
<td>Cooling degree days</td>
</tr>
<tr>
<td>CER</td>
<td>Commission for Energy Regulation</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>CPI</td>
<td>Consumer Price Index</td>
</tr>
<tr>
<td>CPP</td>
<td>Critical peak pricing</td>
</tr>
<tr>
<td>CRB</td>
<td>Commodity Research Bureau</td>
</tr>
<tr>
<td>CT</td>
<td>Causal tree</td>
</tr>
<tr>
<td>DWL</td>
<td>Deadweight loss</td>
</tr>
<tr>
<td>EIA</td>
<td>Energy Information Administration</td>
</tr>
<tr>
<td>EPA</td>
<td>Environmental Protection Agency</td>
</tr>
<tr>
<td>HAC</td>
<td>Heteroskedasticity and autocorrelation consistent</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>HDD</td>
<td>Heating degree days</td>
</tr>
<tr>
<td>H.R.</td>
<td>House Resolution</td>
</tr>
<tr>
<td>IHD</td>
<td>In-home Display</td>
</tr>
<tr>
<td>IP</td>
<td>Initial Production</td>
</tr>
<tr>
<td>ISSDA</td>
<td>Irish Social Science Data Archive</td>
</tr>
<tr>
<td>IV</td>
<td>Instrumental variable</td>
</tr>
<tr>
<td>kW</td>
<td>Kilowatt</td>
</tr>
<tr>
<td>kWh</td>
<td>Kilowatt hour</td>
</tr>
<tr>
<td>LATE</td>
<td>Local average treatment effect</td>
</tr>
<tr>
<td>mcf</td>
<td>Thousand cubic feet</td>
</tr>
<tr>
<td>MMBTU</td>
<td>Million British thermal units</td>
</tr>
<tr>
<td>OLR</td>
<td>Overall load reduction</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary least squares</td>
</tr>
<tr>
<td>RCM</td>
<td>Rubin Causal Model</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root mean square error</td>
</tr>
<tr>
<td>SE</td>
<td>Standard error</td>
</tr>
<tr>
<td>SO$_2$</td>
<td>Sulfur dioxide</td>
</tr>
<tr>
<td>SUTVA</td>
<td>Stable unit treatment value assumption</td>
</tr>
<tr>
<td>TE</td>
<td>Treatment effect</td>
</tr>
<tr>
<td>TOU</td>
<td>Time of use</td>
</tr>
<tr>
<td>TX</td>
<td>Texas</td>
</tr>
<tr>
<td>U.S.</td>
<td>United States</td>
</tr>
<tr>
<td>WTI</td>
<td>West Texas Intermediate</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to express intense gratitude to my committee—Billy Pizer, Richard Newell, Steven Sexton, and Christopher Timmins—for providing guidance and feedback over the course of my graduate studies and beyond. I also appreciate financial support from the Nicholas Institute for Environmental Policy Solutions, the Duke University Energy Initiative, Resources for the Future, and my advisors.
Chapter 1

Introduction

This dissertation includes three papers discussing different aspects of environmental and energy economics and policy. In Chapter 2, “Peaking Interest: How awareness drives the effectiveness of time-of-use electricity pricing”, I analyze what factors drive consumer responses to time-of-use electricity pricing. Such pricing is an increasingly common policy that is designed to reduce reliance on high-cost, inefficient, and often polluting power plants. Using both survey and meter data on Irish households, I find that behavioral factors drive consumer responses: consumer awareness and information are key to inducing responses, and marginal prices have minimal effects.

Chapter 3, “Prices versus Quantities with Policy Updating”,¹ is a theoretical study considering how intertemporal trading (“banking”) under cap-and-trade style policies changes the traditional trade-off between price and quantity regulation (Weitzman 1974). Our theoretical model illustrates an unappreciated advantage of quantity regulation: expected policy updates are incorporated immediately through secondary markets via a no arbitrage condition. This helps us understand the welfare implications of observed price volatility in permit markets and has implications for the design of environmental regulations, such as the debate over the relative merits of a

¹ Note: This work was co-authored with William A. Pizer. Prest’s contribution involved conception of the initial idea, development of the basic modeling framework, derivation of the initial results, and authoring a first draft. In addition, modeling of political “noise”, development of the T-period model, and writing and editing subsequent drafts were undertaken collaboratively.
carbon tax versus “cap and trade” policy.

In Chapter 4, “Trophy Hunting vs. Manufacturing Energy: The Price-Responsiveness of Shale Gas”, we use well-level data to estimate how the shale revolution has changed the price responsiveness of U.S. natural gas supply. We find that U.S. natural gas supply is approximately three times more price responsive as a result of the shale revolution, owing entirely to higher production per well. This improves our understanding of the market dynamics around natural gas supply.

Each of these papers has implications for environmental and energy policy. The first paper aims to understand how to improve the effectiveness of time-of-use electricity pricing policies. The second paper addresses an important feature of climate policy design, given high observed volatility in carbon allowance prices in the United States and Europe. The third paper aims to improve understanding of the recent changes in U.S. natural gas markets, with important implications for the fuel mix in electricity generation (in particular, coal-fired versus gas-fired generation) and hence CO$_2$ emissions.

---

$^2$ Note: this work was co-authored with Richard Newell and Ashley Vissing. Prest’s contribution involved collecting and cleaning the data, developing and estimating econometric models, devising and computing the simulation models, and writing the initial draft. Developing the hazard models and editing the drafts were performed collaboratively.
Chapter 2

Peaking Interest: How awareness drives the effectiveness of time-of-use electricity pricing

2.1 Introduction and Motivation

The marginal cost of electricity production varies greatly by the time of the day. At peak times, delivering another kilowatt hour of energy is expensive and often environmentally unfriendly, whereas off-peak it is low cost. Despite this fact, consumers typically pay a constant price for electricity throughout the day. Recent advances in smart metering technology now allow for charging consumers time-varying prices, encouraging them to reduce their peak consumption or shift it to off-peak periods.

The benefits of dynamic pricing are threefold. First, it can reduce peak load for the grid as a whole, reducing the fixed costs of building extra “peaking” generation capacity to serve load when demand is high. Second, the mismatch between constant retail prices and time-varying marginal costs encourages consumers to over-consume during high-cost peak periods and under-consume during low-cost off-peak periods, resulting in deadweight losses associated with variable generation costs. Third, time-varying pricing can help accommodate the integration of renewable generation into
the grid, thereby reducing dependence on non-renewable sources.\textsuperscript{1} For all of these reasons, it is increasingly important for electricity and environmental policy to understand how consumers respond to time-varying pricing, and what can make these policies more effective.

Economists have long asserted that “getting the prices right” would solve the problem of peak load (Harding and Sexton 2017). However, the existing literature has found that consumer responses to peak pricing have been underwhelming (ibid.). The natural question is why the responses have been modest, and how different policy interventions might increase them. For example, it could be that technology can increase responses, either by providing helpful information to consumers or automating their behavior. Jessoe and Rapson (2014) show that providing consumers with more information about their energy use can amplify the effects of time-varying pricing. Bollinger and Hartmann (2016) and Harding and Lamarche (2016) further show how automating technology, such as thermostats that can be programmed to respond to real-time prices, can have even larger effects.

While information and automation appear to increase households’ responses, the precise causal mechanisms underpinning these findings are often unclear. One idea is a standard neoclassical explanation. For example, it could be that these technologies lower the costs of shifting energy use, and without them households would rationally choose not to change their consumption. Neoclassical explanations also center on the role of price signals, imploring policymakers to “get the prices right” as a way to solve the peak load problem. A related neoclassical idea is that household characteristics moderate responses. For example, perhaps households that own certain types of appliances have lower adjustment costs, or perhaps those with more education find

\textsuperscript{1} For example, in the early evening in California, electricity demand soars as people arrive at home while solar generation declines as the sun sets, resulting in an immediate surge in required generation from fossil fuel plants. By better aligning the timing of demand with that of intermittent renewable supply, dynamic pricing can reduce the costs of meeting environmental goals in the power sector (see, e.g., Finn and Fitzpatrick 2014; Fripp 2016).
it easier to figure out how to respond effectively. These ideas might suggest targeting policies towards households with certain observable characteristics, or perhaps even encouraging households to install certain appliances.

Another theory stems from behavioral economics. Perhaps households are simply inattentive to electricity prices, and providing technology helps makes these prices more salient. Generally, Sexton (2015) suggests that price salience is important for overall electricity consumption, showing evidence that automatic bill pay reduces price salience and increases energy use. More specific to time-varying pricing, Fowlie et al. (2017) show that default effects are important for explaining consumer responses; while this result could possibly be explained by a neoclassical model of switching costs, they argue in favor of more behavioral explanations about price salience.

Despite this variety of potential underlying causal mechanisms, it is often difficult to distinguish between them empirically. This matters because the different mechanisms imply very different lessons for policy design. A price-centric explanation would reinforce the economist’s mantra of “get the prices right,” meaning a policy of simply setting consumer prices equal to the marginal cost of production. Beyond setting correct prices, this theory would suggest a generally laissez-faire approach and generally not recommend additional interventions. A theory emphasizing the role of household characteristics might suggest that policymakers target certain households that are expected to exhibit larger responses (e.g., those with appliances that lower adjustment costs), or perhaps attempt to alter those characteristics (e.g., encouraging different appliances). By contrast, a behavioral theory would suggest focusing on consumer attention and the salience of the policy. In this paper I examine the relative importance of the various competing theories of electricity demand by taking a data-driven approach to investigate the key drivers of household responses to time-varying electricity pricing.
Using data from an time-of-use (TOU) pricing experiment in Ireland, I apply and extend the Athey and Imbens (2016) machine learning algorithm to estimate heterogeneous responses to TOU pricing across a large set of household observables. The experiment involves randomly-assigned pricing schemes (with variation in the magnitude of peak prices) and information provision treatments (including varying billing frequency and providing in-home electricity display), resulting in 16 distinct treatment groups and one control group. The Athey-Imbens algorithm allows me to identify the conditions under which such pricing is more or less effective. This provides a parsimonious way to estimate sources of heterogeneity in treatment effects that are robust to out-of-sample validation, while avoiding multiple testing concerns.

The method allows me to systematically test for heterogeneous responses across a large set of more than 150 household observables. The past literature has not systematically considered this many dimensions (typically focusing on one or two dimensions) due to two major obstacles. First, many past experiments have simply not had access to such rich, high-dimensional data on household observables. Second, even among studies with rich data, testing many dimensions for heterogeneous responses raises methodological concerns about multiple hypothesis testing. The Athey-Imbens algorithm is designed to avoid multiple testing concerns in high-dimensional settings by using hold-out samples to confirm the validity of the estimated effects.

The results indicate that the most important driver of the treatment effect is, by far, consumer awareness. Households who reported in a follow-up survey that they were aware of their time-varying pricing reduced peak consumption by 4.5 times as much as those who were unaware: -10% versus -2.3%. This result suggests the importance of behavioral explanations of consumer behavior. It also highlights the importance of price salience, as previously illustrated by Chetty et al. (2009) and Sexton (2015). More practically, it suggests that substantial additional savings could be achieved by ensuring that households are aware of their tariff structure.
Among households who were aware of their price changes, households with very low baseline energy use did not reduce peak consumption, and may actually increase it in some cases. One explanation for this result is that low-consuming households have less discretionary consumption like television use, so they have fewer opportunities to reduce.

Households receiving more information treatments responded substantially more, consistent with Jessoe and Rapson (2014). The difference is substantial, with an in-home electricity display (IHD) delivering an average treatment effect (-15%) nearly twice as large as a simple bi-monthly energy usage statement (-8%). Information amplified household responses even among those who reported being aware of their pricing mechanism, meaning that this difference does not simply reflect improved awareness. This shows that information provision is crucial when implementing time-of-use pricing, even among those who report being aware of the policy.

Conditional on these sources of heterogeneity (awareness, baseline consumption, and information treatment), I find no other robust relationships between treatment effects and household and house characteristics, despite a long list of such variables in the dataset, such as demographics, house size, appliances owned, and so on. This suggests that, conditional on a household’s baseline consumption, little additional information is necessary to predict treatment effects. In this sense, baseline consumption can be considered a “sufficient statistic” for the larger suite of household observables related to energy use.

Surprisingly, households appear to violate a central law of the economic theory of demand: while they respond to the existence of a price change, they are extremely insensitive to the magnitude of the price change. In particular, point estimates scale less-than-proportionally with the size of the price change (implying a non-constant elasticity), and they are not statistically distinguishable from each other. This effect mirrors results in other contexts where the existence of a financial incentive matters
but the size of the incentive does not. For example, Holladay et al. (2016) find that a subsidy for home energy audits increased uptake, but uptake was no different under a small versus large subsidy. In a non-energy context, Karlan and List (2007) shows that charitable giving increases in response to a matching grant, but larger match ratios do not induce additional giving.

A caveat is that these results are based on an experiment of Irish residential consumers without widespread automation technology. It is possible that other types of consumers (commercial or industrial consumers, or consumers in other countries) may respond differently. It is also possible that households with more automation technology would be more sensitive to the price levels.

While there is significant heterogeneity in the treatment effects during peak periods, there is no detectable heterogeneity during non-peak periods. Neither awareness, baseline energy use, information treatment, nor any other variables predict differential responses during off-peak periods. Households primarily reduced peak use overall, rather than simply shifting peak use to off-peak periods. Nighttime usage increased only modestly (+1.8%), whereas off-peak daytime usage decreased (-2.2%).\(^2\) The latter result may appear somewhat anomalous, as households consumed less despite lower prices. This suggests that adjustment costs across hours of the day lead households to reduce consumption off-peak in anticipation of peak prices.

Given that awareness is the key driver of heterogeneity, I also explore what characteristics of households best predict whether they will be aware of — and hence respond to — the treatment. I find that observable household characteristics have little predictive power for consumer awareness, despite having over one hundred such covariates spanning demographics, appliance ownership, physical house characteristics—

\(^2\) The t-statistics are 1.6 and -2.9, respectively. Overall, the program reduced mean consumption by 2.3\% (\(p = 0.02\)).
tics, and attitudes toward electricity use. For future policies, these results suggest that it is unlikely to be effective to attempt to target policies towards households more likely to be aware.

This study makes two primary contributions. First, it identifies the key sources of treatment effect heterogeneity in time-of-use pricing using new techniques from machine learning, based on Athey and Imbens (2016). The algorithm retains only the dimensions of heterogeneity that are robust to out-of-sample validation, and it eliminates those that are not. This reveals that awareness of the policy is key to program effectiveness, providing direct evidence for the claims in Fowlie et al. (2017). Beyond awareness, only a few factors matter for observable heterogeneity: information and baseline energy consumption. Conditional on these factors, no other observables are robustly related to differential responses, including the degree of the price increase.

This study’s other primary contribution is that it extends the Athey and Imbens (2016) algorithm to multiple treatment groups. This extension allows for using that algorithm to find heterogeneous treatment effects not only across household observables but also across treatment groups. This is important in settings such as this one, where some treatment dimensions (i.e., information) matter for treatment effect heterogeneity while other dimensions (i.e., pricing magnitude) do not.

This paper is organized as follows. In section 2.2 I summarize the relevant literature. In section 2.3, I summarize the program, the data, and treatment/control balance. Section 2.4 establishes the average treatment effect of TOU pricing on consumption by time of the day and discusses the source of identification. In section 2.5, I discuss the estimation of heterogeneous responses. Specifically, section 2.5.1 discusses the Athey-Imbens algorithm and my extensions, and section 2.5.2 presents the estimated heterogeneous responses. Auxiliary results are presented in 2.5.3 and 2.5.4. Finally, section 4.5 concludes.
2.2 Literature

There is a rich literature on the impacts of policies designed to affect residential electricity consumption. First, there are several studies in behavioral economics assessing the effectiveness of “nudges” and social norms on energy consumption. Some studies (Allcott 2011 and Ayres et al. 2013) show that the simple nudge of sending consumers letters comparing their energy usage to that of their neighbors can reduce energy use by as much as 2%. Gans et al. (2013) finds evidence from a natural experiment in Ireland that real-time information alone can substantially reduce electricity consumption. Houde et al. (2013) finds that providing consumers with real-time information through web-based tool reduced energy consumption during mornings and evenings, although they find no heterogeneity in this effect.

Several studies have raised concern that these effects may not always be broadly applicable due to heterogeneity in their impacts. For example Allcott and Rogers (2014) suggests that such effects may decay without sustained information provision. Ito et al. (2015) finds that “moral suasion” — pleas for conservation — had only short-run effects and finds some evidence that it was more effective for higher-income consumers.

The literature has generally found that dynamic electricity pricing can reduce consumption during peak hours, but the effects are rather model (see Harding and Sexton (2017) for a recent survey of the literature). Faruqui and Palmer (2012) shows diminishing returns to larger price increases. More recent studies have supported this point, including Ito et al. (2015) and Carroll et al. (2014). Faruqui and Palmer (2012) also finds that information provision generally enhances the effectiveness of the policies. Wolak (2011) shows that consumers respond to comparable “critical peak pricing” and hourly TOU pricing policies in a similar manner.³ In

³ There are multiple kinds of dynamic electricity pricing. One is “critical peak pricing” (CPP),
more recent evidence, Ito et al. (2015) find that CPP has more persistent effects on peak consumption than behavioral nudges do. Herter and Wayland (2010) also finds significant reductions during CPP events, particularly for larger consumers and those in cool climate zones, but finds that consumers did not respond more to larger price increases. Blonz (2016) finds that commercial and industrial consumers in California respond to CPP pricing, but only in hot inland regions and not in milder coastal regions. In Connecticut, Jessoe and Rapson (2015) finds negligible responses to TOU pricing from commercial and industrial consumers.

This literature often finds that information provision helps increase the effectiveness of dynamic pricing. For example, Jessoe and Rapson (2014) find that consumers given IHDs exhibited stronger and more consistent effects on consumption during CPP events than did consumers without IHDs. Ivanov et al. (2013) similarly find that IHDs amplified the effectiveness of CPP events. Bollinger and Hartmann (2016) also find that IHDs amplified the impacts of dynamic pricing, but they further find that automation of electricity consumption through programmable thermostats has even larger effects. Together, the literature suggests that information provision can help increase the effectiveness of dynamic pricing, and perhaps even more can be done, such as automation or targeting.

Fowlie et al. (2017) discusses the importance of default options in driving electricity consumption in response to TOU pricing. They find strong evidence of a default effect. In particular, they distinguish between households who voluntarily opt in the time-varying pricing (“always takers”) to those who only particulate if time-varying pricing is the default option (so-called “complacent” households). Demand response from “complacent” households is shown to be about half as large as the response

in which consumers face very large price increases (e.g., 300% increases) but only during rare and extreme events (e.g., perhaps few times per year, on very hot days). Another is time-of-use (TOU) pricing under which consumers’ prices routinely vary from hour to hour. To avoid overly complicated pricing schedules, usually only two or three rates are applied (e.g., a low rate at night, a moderate rate during most daytime hours, and a high rate during peak hours).
among “always takers.” They argue that the effect of complacency represents a lack of awareness of the policy, instead of a neoclassical switching cost model. However, they add a caveat that while the results are suggestive that a lack of awareness is behind the smaller responses, the evidence is not dispositive. In this paper, I provide direct support for their argument, showing that consumer awareness is the single most important factor moderating household responses to time-varying pricing.

More generally, while the literature generally finds that time-varying pricing can reduce peak consumption, studies typically do not systematically assess the significant heterogeneity in these effects. A systematic assessment of heterogeneity can help identify the underlying causal mechanisms and the conditions under which pricing policies are most effective, which in turn can improve policy targeting and design. This paper contributes to the literature on time-varying electricity pricing by assessing many possible dimensions of treatment effect heterogeneity, while avoiding multiple hypothesis testing concerns through a modern machine learning algorithm. As a result, I find direct support for arguments that past literature has only suggested.

A recent working paper (Burlig et al. 2017) is relevant because it also applies machine learning to estimating causal effects from high-frequency electricity consumption data. That study uses Lasso-based projections of counterfactual electricity consumption to estimate the impact of energy efficiency investments, finding impacts less than half as large as projected by engineering estimates. While that paper studies a different question, it illustrates the value of applying machine learning techniques to estimate causal effects in electricity interventions.

A working paper by Gillan (2017) corroborates my findings about consumers’ insensitivity to the size of the price change. That study is based on data from California, suggesting that this result is robust, and not simply an artifact of the particular setting of Ireland.
Finally, a small number of studies has explored the data from the same experiment in Ireland, although they have generally focused on different questions than this paper does. Carroll et al. (2014) focus on the extent to which the treatment increased participants’ self-reported knowledge about their energy consumption and how to reduce it; they find that while the program did increase participants’ knowledge, such increases were not correlated with actual reductions in consumption. McCoy and Lyons (2016) find that the treatment actually made participants less likely to install energy-saving technologies, perhaps due to crowding-out effects.

Di Cosmo et al. (2014) estimated average treatment effects from peak pricing and noted relatively minimal differences across tariff levels. Following up on this analysis, the authors investigated heterogeneity in the treatment effect along three dimensions: the education of the chief income earner, their age, and household occupancy status. They found no statistically significant differences in peak responses by education status, and the significance for the other dimensions (age and occupancy) were not reported. No other dimensions of heterogeneity were considered. In a related study, Di Cosmo and O’Hora (2017) revisit whether higher-educated households responded more, finding no significant difference during the first half of the peak period but a significant difference (at the 5% level) in the second half. The magnitudes of these differences were not reported, and no joint test of significance was presented. Pon (2017) finds that the IHD amplified the treatment effect but that this effect decayed over time.

2.3 Data

2.3.1 Description of the Program

The data for this study comes from the smart meter Consumer Behavior Trial (hereafter, the “trial”) run during 2009-2011 by the Commission for Energy Regulation
(CER) in the Republic of Ireland. The experiment was designed to assess the effectiveness of various time-of-use pricing structures and demand-side management stimuli on consumer electricity consumption. In the experiment, CER recruited a nationally representative sample of Irish households. Recruitment was performed via mailer in four waves to ensure the sample was representative on several characteristics including historical electricity consumption. Participation was voluntary, with substantial financial incentives. All households received €25 for each survey completed. In addition, treated households received credits to offset any extra costs from being on a variable tariff. All credits were paid outside of the treatment period (half in the last month of the baseline period, half in the first month after the treatment period) to avoid income effects.

Households were randomly assigned to either treatment and control groups. Smart meters were installed in both treatment and control houses. Before the treatment began, baseline electricity consumption data was collected at the half-hourly resolution from July through December of 2009. The treatment period was January to December 2010. Households who were assigned to be treated were allocated to one of four time-of-use tariffs and also one of four demand-side management stimuli (hereafter “information stimuli”). This four-by-four treatment structure implies sixteen distinct treatment groups.

4 These credits were €30, €50, €70, or €90 to households in the A, B, C, or D tariff groups, respectively. They were paid in two equal installments, one in December 2009 and another in January 2011, each of which is outside of the treatment period.

5 For a detailed summary of the trial, see Commission for Energy Regulation (2011b).

6 The random allocation was designed to ensure balance on covariates across experimental cells. Namely, households were classified into one of several “profiles” based on a principal component analysis of energy usage and survey data. Then, households in each profile were randomly assigned to treatment/control groups. In addition, some participants were moved to different cells after the initial allocation to improve balance between cells.

7 Households were told of their treatment/control assignment late during this benchmark period (December 2009) to avoid potential effects on baseline behavior.

8 There was also 17th treatment group with a relatively small number of participants who faced
The time-of-use pricing component involved higher prices during peak periods (5pm-7pm), significantly lower prices at night (11pm-8am), and slightly lower prices at other times (8am-5pm and 7pm-11pm, referred to as “daytime” hours). All treated households received a refrigerator magnet and sticker explaining the time-of-use pricing scheme. The pricing schedules ranged from somewhat peaked (12 cents per kWh at night vs. 20 cents during peak hours (all values in €)) to very strongly peaked (9 cents vs. 38 cents, respectively). These pricing schedules are illustrated in Figure 2.1. The pricing structures were designed with the goal that the “average” participant who did not change their consumption behavior would not face higher bills on average.

For the information stimuli, all treated households received an energy usage statement along with their bill, containing information about their consumption by time of the day and comparisons to peer households. The first information stimulus group received the energy usage statement along with the household’s bi-monthly bill (that is, every other month). The second information stimulus involved billing consumers more frequently (every month). The third information stimulus billed bi-monthly but provided households with an in-home electricity monitor displaying both real-time and historical information about energy usage and prices. The final stimulus involved providing the bi-monthly billing statement plus an “Overall Load Reduction” (OLR) incentive. The OLR stimulus offered households a €20 bonus if they

---

9 Time-of-use pricing was only active on non-holiday weekdays. For this reason, I only include these days in my analysis, unless otherwise stated. In the appendix, I also conduct a placebo test for a treatment effect during weekends and holidays, finding little to no effect.

10 Peak tariffs were 20, 26, 32, and 38 cents per kWh for tariffs A, B, C, and D, respectively. The corresponding rates during off-peak daytime periods were 14, 13.5, 13, and 12.5 cents. At night, they were 12, 11, 10, and 9 cents. The control group faced a flat price of 14.1 cents for all hours.

11 See appendix for pictures illustrating the monitor, refrigerator magnet/stickers, and energy usage statements.
could reduce their baseline consumption by 10% or more. Control households experienced no change in their electricity rate (a constant 14.1 cents) or billing frequency (bi-monthly).

The program also involved both a pre- and post-trial survey, for which participants received €25 per survey to complete. The surveys involved hundreds of questions covering a wide array of topics, including socio-demographic characteristics, attitudinal questions, physical attributes of the home, and self-reported expectations about home energy use.\(^\text{12}\)

Unfortunately, some households did not complete the surveys. For my analysis, I only include households that did so (74% of the full sample). However, for households that did not, I can still observe their energy consumption and their treatment assignment. This allows me to test whether the average treatment effect is significantly different for those who did not answer the survey. The average effect among the full sample is slightly smaller in magnitude, but not statistically different.\(^\text{13}\) This suggests that the survey response rate is not strongly biasing my treatment effect results.\(^\text{14}\)

\(^{12}\) The pre- and post-trial surveys include 377 questions combined. Full lists of these questions are available from ISSDA, here and here. Not every household was asked every question however. For example, questions about the in-home electricity monitors were not asked of households who did not receive a monitor. Survey questions not answered by all households must be dropped from the analysis, leading to the approximately 150 variables I use. See the appendix for a list of the survey questions included in the analysis.

\(^{13}\) The average effect on peak consumption is 8.3% for the full sample, compared to 8.9% for those who completed both surveys (not significantly different; \(p = 0.38\)).

\(^{14}\) In addition, about 21% of households were lost to attrition, the vast majority of which (83%) was due to a change of supplier. The attrition rate was only slightly higher for control households (25%) than it was for treatment households (20%). An investigation into the reasons for changes in supplier by CER found that “[a]mong attritors who changed supplier [who represent the vast majority of attritors], the reason for the switch appears to have been independent of any potential impact of the Trial. None of these switchers stated that the tariffs and or technology were a factor in this decision. 5% stated that the consumption reports and other information contributed to the decision.” (Commission for Energy Regulation 2011a). I also drop a small number of households because they have extended periods of zero consumption, as these are likely to be frequently uninhabited vacation homes.
The treatment/control distribution of the 3,006 remaining households is shown in Table 2.1.¹⁵

![Figure 2.1: Time-of-Use Pricing Structures](image)

### 2.3.2 Treatment/Control Balance

While treatment/control groups were assigned randomly, the groups did not end up being perfectly balanced. Table 2.2 shows the group means and test statistics for balance for a variety of observable pre-trial characteristics between the control and treatment groups. The top panel of that table includes all variables that were found to be significantly different between the treatment and control groups. The bottom panel shows a selection of the variables with non-significant differences.

¹⁵ Fewer households were assigned to tariffs B and D by design based on CER’s power analysis.
Table 2.1: Treatment and Control Group Assignments

<table>
<thead>
<tr>
<th>Bi-monthly Bill &amp; Energy Usage Statement</th>
<th>Monthly Bill &amp; Energy Usage Statement</th>
<th>Bi-monthly Bill, Energy Usage Statement, and In-Home Display (IHD)</th>
<th>Bi-monthly Bill, Energy Usage Statement, and Overall Load Reduction (OLR) Incentive</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tariff A</td>
<td>195</td>
<td>216</td>
<td>205</td>
<td>216</td>
<td>0</td>
</tr>
<tr>
<td>Tariff B</td>
<td>80</td>
<td>87</td>
<td>72</td>
<td>81</td>
<td>0</td>
</tr>
<tr>
<td>Tariff C</td>
<td>222</td>
<td>217</td>
<td>202</td>
<td>213</td>
<td>0</td>
</tr>
<tr>
<td>Tariff D</td>
<td>80</td>
<td>87</td>
<td>78</td>
<td>77</td>
<td>0</td>
</tr>
<tr>
<td>Control</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>678</td>
</tr>
</tbody>
</table>

| Total                                  | 577                                  | 607                                                          | 557                                                                              | 587     | 678   | 3,006|

Some of the variables that are statistically different between treatment and control groups are directly related to energy consumption. In particular, baseline night-time consumption (variable #9) is larger for treatment households (0.15 kWh per half hour) than for control households (0.14 kWh). Peak and daytime consumption are also somewhat higher, although not statistically different ($p = 0.12$ and $p = 0.07$, respectively).

In total, 122 pre-trial variables are tested, of which 14 (11%) are significantly different for treatment and control at the standard 5% level, although none are significant when using the Bonferroni correction for multiple hypothesis testing. If the groups were perfectly balanced and all variables were independent, we would expect approximately 5% of the variables tested to be significant at this level. One may, then, be concerned that the larger share, 11%, of the variables is significant. However, many of these variables are strongly correlated with each other. For example, variable #3, “Number of Electronics” is equal to the sum of all electronics in the household, including variables #2 (number of large televisions) and #11 (number of desktop computers). Similarly, variable #6, “Number of Residents”, is perfectly collinear with the number of adults (#16) in the home, the number of children (#12), and the indicator variable indicating whether any children live in the home (#5).
These variables are perfectly collinear, so interpreting them as independent sources of imbalance would be inappropriate.

To adjust for this perfect collinearity, I also estimate a linear probability model of treatment status on pre-treatment household observables. The results of this estimation are shown in Table 2.3. Here, the observables do little to predict treatment status, as only 2% of the variables tested are statistically significant at the 5% level, and again none are significant with the Bonferroni correction. In particular, the coefficients on the baseline consumption variables are small and insignificant, suggesting that energy consumption is balanced conditional on household observables. This suggests the treatment and control groups are balanced on average over the full set of observables, if not on each specific one taken individually.

While the small imbalances between the groups are not generally statistically significant, it is worth considering differences on a key variable of interest: average peak electricity consumption during the baseline period (variable #15 in Table 2.2). Baseline peak consumption is slightly larger for the treatment group (0.44 kWh per half hour) than for the control group (0.42 kWh). This amounts to an approximately 5% difference. While the difference is not statistically significant, its magnitude is non-trivial relative to typical treatment effect sizes in the literature. For this reason, I pursue a difference-in-differences strategy to net out time-invariant factors that are different between the treatment and control groups. Given the differences in treatment in control baseline consumption, a single treatment/control difference

---

16 A logistic regression produces the exact same set of statistically significant coefficients as the linear probability model and hence has the same interpretation. I present the linear probability model because it is easier for the reader to interpret the coefficients.

17 Further, performing model selection by Lasso concludes that all or nearly all covariates are unrelated to treatment, depending on the criteria used to choose the tuning parameter. As discussed at length in section 2.5.4 below, there are two standard metrics for choosing tuning parameter in the Lasso: the RMSE-minimizing lambda and the 1 standard error lambda. Here, the conservative 1 standard error rule finds that no covariates are related to treatment status, and the RMSE-minimizing lambda finds that only the number of televisions and social class variables are related.
would be significantly biased.\textsuperscript{18} This highlights the importance of the difference-in-differences approach.\textsuperscript{19}

2.4 Average Treatment Effects Using Difference-in-Differences

In this section, I use a difference-in-differences approach to estimate average treatment effects and highlight the sources of identification. In the subsequent section, I explore the heterogeneity in these effects.

2.4.1 Half-Hourly Average Treatment Effects

I estimate the average treatment effect for each 30-minute interval of the day in a difference-in-differences framework. I do not distinguish between the sixteen different treatment groups for the moment, as I will estimate that dimension of heterogeneity in subsequent sections. For each of the 48 half-hour periods of the day, I separately estimate the following specification for each half hour of the day $h \in \{1, 2, ..., 47, 48\}$:

$$\ln(Y_{i,h,t}) = \beta_h W_{i,t} + \alpha_{i,h} + \lambda_{w,h} + \epsilon_{i,h,t}. \hspace{1cm} (2.1)$$

$Y_{i,h,t}$ is the electricity consumption by household $i$ in during time window $t$ (a half-hour, e.g., 5:00pm-5:30pm on March 9, 2010). $W_{i,t}$ is an indicator equal to one for treatment households during the treatment period, and is zero otherwise.\textsuperscript{20} $\alpha_{i,h}$ and

\textsuperscript{18} For example, if the treatment effect were actually -7%, then estimating the treatment effect by subtracting treatment and control means would recover a severe underestimate of -2% ($=-7\%$ treatment effect + 5% bias).

\textsuperscript{19} Perfect balance is not required for the Athey and Imbens (2016) estimator I employ in the subsequent section. That estimator instead requires only the weaker assumption of unconfoundedness (i.e., that conditional on observables treatment is uncorrelated with the outcome variable). Further, I show in the appendix that my results are robust to using propensity score weighting to correct for imbalance on observables.

\textsuperscript{20} Formally, $W_{i,t}$ is defined as follows:

$$W_{i,t} = \begin{cases} 1, & \text{if } i \text{ is treated and } t \text{ is in treatment period} \\ 0, & \text{otherwise} \end{cases}$$
<table>
<thead>
<tr>
<th>Variable</th>
<th>Control Mean</th>
<th>Treatment Mean</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unbalanced Variables (p &lt; 0.05)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Employment status: Retired (Indicator)</td>
<td>0.38</td>
<td>0.31</td>
<td>3.40</td>
<td>0.001</td>
</tr>
<tr>
<td>2. Number of Large Televisions (21+ inch)</td>
<td>1.19</td>
<td>1.31</td>
<td>-3.36</td>
<td>0.001</td>
</tr>
<tr>
<td>3. Number of Electronics</td>
<td>3.74</td>
<td>4.04</td>
<td>-3.03</td>
<td>0.003</td>
</tr>
<tr>
<td>4. Age Group: 65+ (Indicator)</td>
<td>0.28</td>
<td>0.23</td>
<td>2.79</td>
<td>0.01</td>
</tr>
<tr>
<td>5. Has Children Under 15 in Home (Indicator)</td>
<td>0.23</td>
<td>0.28</td>
<td>-2.78</td>
<td>0.01</td>
</tr>
<tr>
<td>6. Number of Residents</td>
<td>2.60</td>
<td>2.76</td>
<td>-2.65</td>
<td>0.01</td>
</tr>
<tr>
<td>7. Social Class: AB, Manager/Professional (Indicator)</td>
<td>0.12</td>
<td>0.15</td>
<td>-2.58</td>
<td>0.01</td>
</tr>
<tr>
<td>8. Education: Primary (Indicator)</td>
<td>0.15</td>
<td>0.11</td>
<td>2.51</td>
<td>0.01</td>
</tr>
<tr>
<td>9. Baseline Average Consumption (Night Hours)</td>
<td>0.14</td>
<td>0.15</td>
<td>-2.42</td>
<td>0.02</td>
</tr>
<tr>
<td>10. Internet Access in Home (Indicator)</td>
<td>0.66</td>
<td>0.71</td>
<td>-2.24</td>
<td>0.02</td>
</tr>
<tr>
<td>11. Number of Desktop Computers</td>
<td>0.48</td>
<td>0.53</td>
<td>-2.18</td>
<td>0.03</td>
</tr>
<tr>
<td>12. Number of Children Under 15 in Home</td>
<td>0.43</td>
<td>0.52</td>
<td>-2.11</td>
<td>0.04</td>
</tr>
<tr>
<td>13. Housing Status: Own with Mortgage (Indicator)</td>
<td>0.35</td>
<td>0.40</td>
<td>-2.09</td>
<td>0.04</td>
</tr>
<tr>
<td>14. Others in Household Use Internet Regularly (Indicator)</td>
<td>0.53</td>
<td>0.57</td>
<td>-2.01</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Selected Balanced Variables (p ≥ 0.05)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Baseline Average Consumption (Peak Hours)</td>
<td>0.42</td>
<td>0.44</td>
<td>-1.85</td>
<td>0.07</td>
</tr>
<tr>
<td>16. Number of Adults in Home</td>
<td>2.16</td>
<td>2.24</td>
<td>-1.74</td>
<td>0.08</td>
</tr>
<tr>
<td>17. Cook stove type: Electric (Indicator)</td>
<td>0.72</td>
<td>0.69</td>
<td>1.62</td>
<td>0.10</td>
</tr>
<tr>
<td>18. Number of Laptop Computers</td>
<td>0.65</td>
<td>0.71</td>
<td>-1.61</td>
<td>0.11</td>
</tr>
<tr>
<td>19. Baseline Average Consumption (Day Hours)</td>
<td>0.29</td>
<td>0.30</td>
<td>-1.56</td>
<td>0.12</td>
</tr>
<tr>
<td>20. Unemployed, not seeking job (Indicator)</td>
<td>0.03</td>
<td>0.04</td>
<td>-1.52</td>
<td>0.13</td>
</tr>
<tr>
<td>21. Home Heat: Solid Fuel (Indicator)</td>
<td>0.29</td>
<td>0.26</td>
<td>1.46</td>
<td>0.14</td>
</tr>
<tr>
<td>22. Interested in changing energy use for environment*</td>
<td>1.38</td>
<td>1.34</td>
<td>1.41</td>
<td>0.16</td>
</tr>
<tr>
<td>23. Female (Indicator)</td>
<td>0.47</td>
<td>0.50</td>
<td>-1.02</td>
<td>0.31</td>
</tr>
<tr>
<td>24. Education: Secondary to Certificate (Indicator)</td>
<td>0.16</td>
<td>0.17</td>
<td>-0.86</td>
<td>0.39</td>
</tr>
<tr>
<td>25. Satisfied with billing frequency*</td>
<td>2.84</td>
<td>2.86</td>
<td>-0.47</td>
<td>0.64</td>
</tr>
<tr>
<td>26. Expect to Choose More Efficient Appliances*</td>
<td>1.34</td>
<td>1.35</td>
<td>-0.35</td>
<td>0.73</td>
</tr>
<tr>
<td>27. Number of Immersion Water Heaters</td>
<td>0.77</td>
<td>0.77</td>
<td>-0.15</td>
<td>0.88</td>
</tr>
<tr>
<td>28. Home Style: Terraced (Indicator)</td>
<td>0.14</td>
<td>0.14</td>
<td>-0.14</td>
<td>0.89</td>
</tr>
<tr>
<td>29. Number of Washing Machines</td>
<td>0.99</td>
<td>0.99</td>
<td>-0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>30. Unemployed, seeking job (Indicator)</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Observations: 3,006
Number of Variables Tested: 122
Number of Variables Not Shown: 92
Number of Variables Significant (5% level): 14
Share of Variables Significant (5% level): 11.5%

Notes: Due to space limitations, only a subset of the 122 tested variables are shown. The suppressed variables are all statistically insignificant at the 5% level. The full set of t-tests is available upon request. The variables are presented in ascending order of p-value. Baseline Average Consumption variables units are kWh per 30 minute interval. This table was generated using the stargazer package (Hlavac 2015) for R.
* These variables featured numeric responses, where respondents reported on a 1-5 scale to what extent they agree (1) or disagree (5) with the statement.
Table 2.3: Treatment/Control Balance: Linear Probability Model of Treatment on Covariates

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Treated (Indicator)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Average Consumption (Peak Hours)</td>
<td>0.03 (0.07)</td>
</tr>
<tr>
<td>Baseline Average Consumption (Night Hours)</td>
<td>0.14 (0.15)</td>
</tr>
<tr>
<td>Baseline Average Consumption (Day Hours)</td>
<td>0.09 (0.12)</td>
</tr>
<tr>
<td>Employment Status: Retired (Indicator)</td>
<td>0.04 (0.05)</td>
</tr>
<tr>
<td>Number of Large Televisions (21+ inch)</td>
<td>0.02** (0.01)</td>
</tr>
<tr>
<td>Age Group: 65+ (Indicator)</td>
<td>0.13 (0.11)</td>
</tr>
<tr>
<td>Has Children Under 15 in Home (Indicator)</td>
<td>0.06 (0.04)</td>
</tr>
<tr>
<td>Number of Adults in Home</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Social Class: AB, Manager/Professional (Indicator)</td>
<td>0.03 (0.08)</td>
</tr>
<tr>
<td>Education: Primary (Indicator)</td>
<td>0.07 (0.04)</td>
</tr>
<tr>
<td>Internet Access in Home (Indicator)</td>
<td>0.01 (0.02)</td>
</tr>
<tr>
<td>Number of Desktop Computers</td>
<td>0.01 (0.02)</td>
</tr>
<tr>
<td>Number of Children Under 15 in Home</td>
<td>0.01 (0.02)</td>
</tr>
<tr>
<td>Others in Household Use Internet Regularly (Indicator)</td>
<td>0.003 (0.02)</td>
</tr>
<tr>
<td>Cook stove type: Electric (Indicator)</td>
<td>0.13** (0.06)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.43 (0.28)</td>
</tr>
<tr>
<td>Observations</td>
<td>3,006</td>
</tr>
<tr>
<td>R²</td>
<td>0.03</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>-0.01</td>
</tr>
<tr>
<td>F Statistic</td>
<td>0.76</td>
</tr>
</tbody>
</table>

| Number of Covariates | 109 |
| Number of Covariates Not Shown | 94 |
| Number of Covariates Significant (5% level) | 2 |
| Share of Covariates Significant (5% level) | 1.8% |

*p<0.1; **p<0.05; ***p<0.01

Notes: Standard errors shown in parentheses. Due to space limitations, only a subset of the 109 included covariates are shown in this table. The full results are available upon request. All suppressed covariates have statistically insignificant coefficients (at the 5% significance level). There are fewer covariates in this table than in the t-tests in Table 2.2 because of perfect collinearity (e.g., number of residents equals the sum of number of adults and children). Baseline Average Consumption variables units are kWh per 30 minute interval. This table was generated using the stargazer package (Hlavac 2015) for R.
\(\lambda_{w,h}\) are household and week-of-sample fixed effects.\(^{21}\) The parameters of interest are \(\beta_h\), particularly during the peak periods, when prices are highest. These represent the average treatment effect for each hour of the day.\(^{22}\) The point estimates for \(\beta_h\), along with 95% confidence intervals, are shown in graphical form in Figure 2.2.

The results indicate a substantial reduction in energy consumption during peak hours, consistent with the substantially higher peak-period price (as shown in Figure 2.1).\(^{23}\) Similarly, low prices at night led to moderate increases in electricity consumption at the beginning of the night time price period (11pm-12:30am), but this effect largely disappears for the remainder of the night hours.

During non-peak daytime hours, there appears to be a moderate reduction in energy usage during some hours, although this effect is only statistically significant for 5 of the 26 half-hour-long periods. This is interesting however because time-of-use prices during these hours were lower than the prices both for the control group and for the treatment group before the trial was implemented (see Figure 2.1). Taken literally, this would imply an upward-sloping demand curve.

One explanation for this phenomenon is that adjustment costs prevent households from changing their consumption on an hourly basis during workdays. Under this theory, households know that peak prices will be charged late in the day, but they will not necessarily be at home at 5:00pm to turn off devices. Anticipating this, they

\(^{21}\) I do not need to include treatment assignment and treatment period indicators as separate regressors because they are captured by the household and week-of-sample fixed effects, respectively.

\(^{22}\) Unless otherwise stated, throughout the paper I am addressing local average treatment effects. As previously mentioned, compliance was imperfect due to attrition, meaning all my estimates must be considered LATEs. However, Commission for Energy Regulation (2011b) suggests attrition was largely unrelated to treatment status.

\(^{23}\) In the appendix section A.2.3, I estimate whether the household responses to peak pricing vary over time. The responses could theoretically grow or shrink over time. Responses might grow over time if households take time to invest in energy-saving devices, or they could shrink due to declining interest and attention to peak prices. In fact, I find that there is no strong temporal trend in household responses to peak prices.
adjust appliances before leaving home in the morning. This is consistent with the first significant reduction in usage occurring around 9:00am and persisting.\(^{24}\)

Another explanation is that the information treatments had an independent effect on daytime consumption that more than offsets the consumption-encouraging effects of lower daytime prices. However, this explanation is inconsistent with the placebo test presented in the appendix, which showed negligible effects on weekends and holidays, when TOU pricing was not in effect but the information treatments would still be salient. A final explanation is that consumers respond to hour-to-hour *changes* in prices, rather than price levels. For example, consumption falls somewhat at 8am-10am when prices increase from the night rate (very low) to the day rate (somewhat low). Similarly, consumption increases at 11pm when prices decrease from the day rate to night rate.

A simple way to observe what is identifying the treatment effects is to consider the simple average daily consumption profiles by treatment and control groups, both during the baseline period and during the treatment period. This is shown in Figure 2.3.

Two important features of this figure stand out. First, the consumption profile for the control group does not significantly change from baseline to treatment period. This suggests that the treatment period (January to December 2010) featured nearly identical average demand conditions as the baseline period (July to December 2009).\(^{25}\) This reveals that the identification of the treatment effect is driven primarily by within-household changes in consumption patterns before and after the treatment period. In particular, a single difference of baseline and treatment period

\(^{24}\) A similar effect was found by Blonz (2016) for commercial and industrial users. Jessoe et al. (2014) also found anomalous responses to a price change, where households reduced electricity consumption in response to a price decrease.

\(^{25}\) Note further that the consumption profiles look very similar, even though the baseline and treatment periods did not cover the same calendar year. The results are robust to only using the second half of the treatment period in order to align the calendar months across periods.
consumption for the treatment group (rather than difference-in-differences) yields almost exactly the same average treatment effect.

The lack of any strong time trend for the control group supports the parallel trends assumption required for the difference-in-differences estimator. For the parallel trend assumption to be violated, the treatment group would have had to exhibit a particular time trend absent treatment, even though the control group evidently did not. In the appendix, I further assess pre-trends, finding that they are similar for the treatment and control groups. The other important assumption for difference-in-
differences is the stable unit treatment value assumption (SUTVA), which requires that control households are not affected by treatment households (or visa versa). This is plausible because this was a nationwide program of only a few thousand geographically-dispersed households, so participating households are unlikely to have interacted on any significant scale.

The second evident feature is that treatment households do not ramp up their demand as rapidly during the treatment period as they did during the baseline period. This is the primary source of identification for the average treatment effect. For households assigned to treatment, consumption peaked at around 0.46 kWh\textsuperscript{26} during the baseline period compared to about 0.407 kWh during the treatment period, a reduction of approximately 11\%.\textsuperscript{27}

As previously mentioned, the treatment group appears to have somewhat higher baseline consumption in peak periods than the control group does. This difference is not quite statistically significant at standard levels ($p = 0.07$). However, the difference is non-trivial in magnitude. As a result, a naive difference between the treatment and control groups would understate the treatment effect during peak hours by a factor of two, and it would estimate the wrong sign for the treatment effect during daytime hours. This highlights the importance of a difference-in-differences approach.

\textsuperscript{26} All kWh figures are in units of energy consumed during the half hour interval. To convert to load in kW, multiply by 2. E.g., 0.46 kWh consumed during 30 minutes is equivalent to a rate of 0.92 kWh consumed per hour, or a load of 0.92 kW.

\textsuperscript{27} This 11\% reduction is slightly larger than the treatment effects shown in Figure 2.2 for four reasons: rounding, it is a difference rather than a difference-in-differences, the log approximation in equation (2.1), and Jensen’s inequality (that is, the average of the log is not the same as the log of the average).
2.5 Heterogeneity in Treatment Effects

As shown in the previous section, the program led to substantial treatment effects during peak hours, and there is some evidence for smaller effects during off-peak hours. In this section, I analyze the heterogeneity (on observables and across treatment groups) in the treatment effects.

2.5.1 Method

Overview of Athey-Imbens Causal Tree Algorithm

I use the Athey and Imbens (2016) causal tree (CT) algorithm which uses the algorithm uses regression trees and cross-validation techniques to estimate heterogeneous
treatment effects. In this section, I provide an overview these methods. Readers already familiar with these models may wish to skim it or skip to the next subsection.

The CT estimator uses the Rubin Causal Model (RCM) as a framework. The RCM postulates two potential outcomes, $Y_i(1)$ and $Y_i(0)$ (e.g., electricity consumption), for each unit $i$ (e.g., household): one where the household is treated ($W_i = 1$) and another in the case where the household is not treated ($W_i = 0$). Since a household cannot be both treated and untreated simultaneously, the econometrician observes only one of these two potential outcomes:

$$Y_{i \text{obs}} = Y_i(W_i) = \begin{cases} Y_i(1) & \text{if } W_i = 1 \\ Y_i(0) & \text{if } W_i = 0 \end{cases}$$

(2.2)

The true treatment effect is then $Y_i(1) - Y_i(0)$. Much of the applied causal inference literature involves estimating the average of this treatment effect (LATE), denoted $\tau$:

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)].$$

(2.3)

This typically is estimated using the sample analogue,

$$\hat{\tau} \equiv \frac{1}{n_T} \sum_{i \in S_T} Y_i - \frac{1}{n_C} \sum_{i \in S_C} Y_i.$$  

(2.4)

where $S_T$ and $S_C$ represent the treatment and control groups, respectively, and $n_T = |S_T|$ and $n_C = |S_C|$ are the number of households in the treatment and control groups, respectively.\(^{28}\)

Whereas this represents the overall average treatment effect on treated households (the local average treatment effect, or LATE), the Athey and Imbens (2016) consistency of this estimator requires assumptions on balance, propensity score weighting, a difference-in-differences technique, or other similar method. I ignore this distinction here for expositional purposes, but I use the difference-in-differences estimator to ensure consistency, followed by robustness checks using propensity-score weighting.\(^{28}\)
algorithm (hereafter referred to as the causal tree (CT) algorithm) focuses on estimating heterogeneity in this effect. Specifically, it estimates conditional average treatment effects (CATE) — that is, the treatment effect conditional on a particular value of observables ($X_i$). The CATE is defined as the treatment effect as a function of $x$,

$$\tau(x) = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = x]. \quad (2.5)$$

An example of a conditional average treatment effect is the average effect among college-educated women over the age of 65. With many different observables $X_i$, some of which may be continuous or interact with each other, estimating $\tau(x)$ is a daunting task, particularly if one wishes to estimate it non-parametrically.

The CT algorithm uses regression trees to estimate $\tau(x)$ in a parsimonious yet non-parametric manner. In particular, the CT algorithm recursively splits the data along covariates into hyper-rectangles, estimating the treatment effect separately within each subset, called “leaves.” This is called growing the tree. Setting aside how the tree is grown, once the splits are found, estimating the CATEs ($\tau(x)$) is straightforward. Given a tree (denoted $\Pi$), the CATE for an individual with characteristics $x$ is simply the difference in conditional means for treatment and control outcomes falling in the same leaf as $x$ (denoted $\ell(x, \Pi)$):

$$\hat{\tau}(x) = \hat{\mu}(x; S_T, \Pi) - \hat{\mu}(x; S_C, \Pi), \quad (2.6)$$

where $\hat{\mu}(x; S, \Pi)$ is simply the conditional mean of observed outcomes in group $S$ (i.e., treatment or control) in the leaf of tree $\Pi$ where $x$ falls.

$$\hat{\mu}(x; S, \Pi) = \frac{1}{|\{i \in S : X_i \in \ell(x, \Pi)\}|} \sum_{i \in S: X_i \in \ell(x, \Pi)} Y_i. \quad (2.7)$$

\[29\] This paper presumes that the reader has some understanding of regression trees. See the appendix for a brief explanation of regression trees.
This is analogous to simply computing the LATE for a each subgroup in the data, where the tree determines the subgroups.

The key innovations in Athey and Imbens (2016) are that it solves two practical difficulties in growing the tree. The first is how to grow the tree to begin with, which is far simpler for standard regression trees that are meant to predict a single outcome variable, $Y_i$, than for treatment effects $\tau$. The second is how to determine the optimal size of the tree, since cross-validation techniques require the researcher to know the ground truth in the outcome of interest. In standard regression problems, the outcome of interest is observed (the outcome, $Y_i$), but in causal inference it is not (the treatment effect, $\tau$).

Athey and Imbens (2016) solves these problems in multiple ways, but the simplest way is through a simple transformation of the outcome variable, $Y$. Denoting the overall probability of treatment $p = n_T/(n_T + n_C) = n_T/n$, they transform the outcome variable as follows:

$$Y_i^* = Y_i \frac{W_i - p}{p(1 - p)} = \begin{cases} 
Y_i/p & \text{if } W_i = 1 \\
-Y_i/(1 - p) & \text{if } W_i = 0
\end{cases}$$

(2.8)

As a result of this transformation, the simple average of $Y_i^*$ is equal to the LATE:

$$\frac{1}{n} \sum_i Y_i^* = \frac{1}{n} \left[ \sum_{i \in S_T} \frac{Y_i}{p} + \sum_{i \in S_C} \frac{-Y_i}{1 - p} \right]$$

$$= \frac{1}{n} \left[ \sum_{i \in S_T} \frac{Y_i}{n_T/n} + \sum_{i \in S_C} \frac{-Y_i}{n_C/n} \right]$$

$$= \frac{1}{n_T} \sum_{i \in S_T} Y_i - \frac{1}{n_C} \sum_{i \in S_C} Y_i$$

$$= \hat{\tau}$$

(2.9)

In this way, the average of the transformed variable captures the average treatment effect, without any explicit need to further distinguish between treatment and con-
trol groups. This equality holds exactly for the full sample, and it also holds in expectation for subsets of the data after conditioning on observables.\footnote{It holds in expectation for each conditional subset of the data, but may not hold exactly for any particular subset. This is because the share of observations treated in that subset of the data may not be exactly the same as the share treated in the full sample due to sampling variability. However, in a randomized experiment, the expected probability of treatment for a subset equals the overall treatment probability, which is why the equality holds in expectation.} Therefore, one can then simply use standard tree-based regression methods and cross-validation approaches to predict the conditional mean of this transformed variable $Y_i^*$, in place of the actual outcome variable.\footnote{While Athey and Imbens (2016) note a number of problems with this simple approach, their algorithm is somewhat more sophisticated than this example presents in order to correct for these. For example, to the extent that treatment probability for a particular leaf differs from the overall probability $p$, the average of $Y_i^*$ will differ somewhat from the actual CATE among observations in that leaf. To solve that, one simply computes the actual CATE for those observations. Another problem is that a tree grown using $Y_i^*$ could potentially form leaves with no control observations. That is solved by precluding splits that would result in branches with fewer than a pre-defined number of treatment and control observations. See Athey and Imbens (2016) for a full description of the many variations on their algorithm. I use a form of their transformed outcome tree to grow the tree, but then report the true CATEs in the nodes. I use this form because simulations suggest that the alternative forms have difficulty finding some CATEs in the presence of certain kinds of correlations in the data generating process.}

While the Athey and Imbens (2016) algorithm is somewhat more sophisticated than this, it illustrates how one can translate a difficult treatment effect heterogeneity problem into a somewhat more manageable non-parametric regression problem. I use the more sophisticated version of the algorithm and add two extensions, described in the next subsections.\footnote{An even more sophisticated algorithm is causal forests, which further extends causal trees to use random forests. As a sensitivity, I also conducted this analysis. The results were very similar. I use the simpler causal tree but causal forests are much more difficult to interpret, as that method produces thousands of individual treatment effects that are not easily summarized in a table or graphic.}

### Extensions to the Athey-Imbens Causal Tree Algorithm

The CT algorithm is designed for standard experiments with a single treatment group and a single observation per treated unit. I extend it to suit the circumstances of this data, which features multiple treatment groups and calls for a difference-in-
First, I extend this algorithm to a difference-in-differences framework where households are observed both before and after treatment. This extension is straightforward, as one need only replace the outcome variable $Y_i$ in the CT algorithm with its change from pre- to post-treatment means, $\Delta Y_i = \bar{Y}_{i,t'} - \bar{Y}_{i,t}$, where the mean is taken over time for each period, $t'$ (the treatment period) and $t$ (baseline period). This captures the first difference, and indeed has better statistical properties than panel data methods, as documented in Bertrand et al. (2004). The CT algorithm then computes the difference in these differences.\(^{33}\)

Second, I extend the CT algorithm to estimate heterogeneity in treatment effects across multiple treatment groups. This extension is more substantial. The natural way to implement multiple treatment effects in the CT algorithm would be to include an indicator variable for each of the $m \in \{1, \ldots, M\}$ treatment groups, denoted $W_{i,m}$, allowing the algorithm to find heterogeneity in treatment effects across those indicators. For example, if treatment group $m$ resulted in larger treatment effects than others, the tree could form a new branch for all observations with $W_{i,m} = 1$. The practical problem with this approach is that such a split would result in no control observations in that side of the branch, making it impossible to estimate treatment effects.

The solution to this is to replace each control observation with $M$ copies of it, each pseudo-assigned to a different treatment group. That is, for each control observation

\(^{33}\) If households are observed the same number of times before and after the treatment, the resulting estimates of the LATE and CATEs are numerically equivalent to the standard difference-in-differences estimates.
i and treatment $m \in \{1, \ldots, M\}$, generate a new pseudo-observation $i_m$ with

\[
Y_{i_m} = Y_i \\
X_{i_m} = X_i \\
W_{i_m} = W_i (= 0)
\]

For $m' \in \{1, \ldots, M\}$

\[
W_{i_{m'}} = \begin{cases} 
1 & \text{for } m' = m \\
0 & \text{for } m' \neq m,
\end{cases}
\]

for a total of $M \times n_C$ control observations, in place of the original $n_C$ observations. This leads to a new pseudo-sample size of $\tilde{n} = (n_T + M n_C)$. With this transformed dataset, the tree algorithm can split on the treatment group covariates ($W_{i_{m'}}$) while continuing to use $W_i$ as the treatment indicator. Meanwhile, duplicating control group observations does not alter the means of control outcomes or observables. As a result, the LATE and CATE estimates on the transformed data are numerically equivalent at every branch to the corresponding estimates on the untransformed data.

To see this formally, consider the mother node, before any splits have occurred, where there is no difference between the LATE and the CATE. Simply considering the LATE, it is easy to show that the LATE of the transformed data is numerically equivalent to the LATE of the untransformed data. The estimated LATE on the transformed dataset is (denoting the set of the transformed control observations $\tilde{S}_C$)

\[
\frac{1}{n_T} \sum_{i \in S_T} Y_i - \frac{1}{M \times n_C} \sum_{i \in \tilde{S}_C} Y_i = \frac{1}{n_T} \sum_{i \in S_T} Y_i - \frac{1}{M \times n_C} M \sum_{i \in \tilde{S}_C} Y_i
\]

\[
= \frac{1}{n_T} \sum_{i \in S_T} Y_i - \frac{1}{n_C} \sum_{i \in \tilde{S}_C} Y_i
\]

\[
= \hat{\tau}, \quad (2.10)
\]

which is the same as the LATE estimate of the untransformed data.
Moreover, the CATE of the transformed data at any node of the tree is also equal to the CATE for the original dataset, even after splitting on any combination of observables and/or treatment groups. Showing this involves only two differences compared to the above logic. First, the means are conditional on $X_i$ and $W_i^m$. And second, the values of $M$ in the second term of equation (2.10) are decremented by one for each treatment group that has previously been split upon, as those groups have been previously diverted into another branch. This latter difference does not affect the estimated CATE, because the decrement appears in both the numerator and the denominator of the second term, and therefore cancels.

When there are overlapping dimensions of treatment, the process can be simplified somewhat to reduce computation time. One need not create an additional observation and indicator variable for each permutation of the different treatment dimensions. For example, in my data, households received one of four information treatments and one of four pricing schedules, resulting in 16 distinct treatment groups. In this case, the process outlined above can be executed twice, with $M = 4$ each time, leading to $8 (= 4 + 4)$ copies of the control observations, rather than $16 (= 4 \times 4)$, resulting in a pseudo-sample size of $\tilde{n} = (n_T + (4 + 4)n_c)$.

While the point estimates of the transformed data are equivalent to those of the original data, the duplicated observations are perfectly correlated (i.e., the observations are not independent) and threaten to bias standard errors. To correct for this, one can either cluster standard errors at the household level or compute standard errors using the original dataset with standard methods. A related issue arises in cross validation because now households appear in the data multiple times, and standard resampling methods are likely to assign a single household’s multiple pseudo-observations to separate CV groups. To solve this, households should be block-assigned to CV groups.
2.5.2 Causal Tree Estimation Results

Results

I estimate the causal tree algorithm using difference-in-differences with multiple treatment groups as described in the previous section. In practice, this means that the outcome variable for household $i$ is the percentage change in $i$’s average peak consumption between the baseline and treatment period. In particular, this can be interpreted as estimating an equation of the following form:

$$\Delta Y_i = \mu + \tau W_i + \varepsilon_i$$  \hspace{1cm} (2.11)

where $\Delta Y_i$ is percentage change in household $i$’s average peak-period consumption between the baseline and treatment period, and as before $W_i$ is an indicator variable for treatment. This is a difference-in-differences estimate because $\Delta Y_i$ is the first difference (before/after for each household), and $\tau$ represents the mean difference in these differences between the treatment and control groups.\(^{34}\) Hence, $\tau$ is the difference-in-differences estimator for the average treatment effect. The causal tree algorithm determines the subgroups of the data for which to estimate the $\tau$ separately, giving the treatment effect as a function of covariates, or $\tau(x)$.

Figure 2.4 shows the results from the causal tree estimation for peak consumption. Each box in that figure represents a “node” of the tree, corresponding to the

\(^{34}\) This differs from the panel-data-style difference-in-differences estimator, which involves method involves regressing $Y_{i,t}$ (as opposed to its change over time) on indicators representing both treatment group and treatment period. In order to apply the causal tree algorithm, I must difference baseline and treatment consumption to obtain a single outcome per household. This is because the causal tree algorithm is designed for cross-sectional data, not panel data. Estimating the treatment effects using the panel method produces nearly identical estimates, differing only slightly due to the log approximation error and the different number of hours in the baseline and treatment periods. While one may think that using the larger sample size of a panel dataset would produce smaller standard errors, in these data the resulting standard errors are approximately the same in both methods after clustering. This is because clustering accounts for the very high within-household correlation in energy consumption over time. See Bertrand et al. (2004) for evidence that aggregating extended time-series into a small number of observations (as I do here) can actually improve the accuracy of standard errors in difference-in-differences estimators.
particular subset of the samples satisfying all conditions from the branches above it. Each box shows the estimated treatment effect (TE), its standard error (SE), and the number of treated observations in that node. The colors of the nodes correspond to the size of the treatment, with green corresponding to larger reductions in peak consumption.

The top node (node [1]) shows the local average treatment effect (LATE) for the full dataset of -8.9% of baseline peak consumption. The first split in the data found by the CT algorithm is awareness: households that reported being aware of the change in their tariff structure exhibit a 4.5-times larger response (-10.3%, node [2]) compared to those who were not (-2.3%, node [11], not statistically significant).

One may be concerned that awareness is endogenous, raising a problem of finding a comparable subset of the control group who would be unaware (or aware) had they been treated. Several factors allay this concern. First, the key identifying variation derives from the within-household change in peak consumption between the treatment period and baseline period (see Figure 2.3), instead of the differences between the treatment and control groups. In particular, the estimated treatment effects are approximately unchanged if I drop control group entirely, which avoids the need to find comparable control groups.

Second, as shown in section 2.5.4, the aware and unaware groups are strongly balanced on nearly all other observables, meaning finding a comparable subset of the control group would have little effect on the results.

35 The standard errors are computed using an OLS regression of equation (2.11) for each relevant subgroup. I do not report p-values for these estimates. Using standard tools of inference based on the standard errors will be biased, likely rejecting the null too often, because the model was both chosen and estimated using the data. The method for computing p-values after model selection methods such as this is not yet settled in the literature.

36 This alternative relies only on the variation of baseline- versus treatment-period, rather than the difference-in-differences estimator which also uses the control versus treatment group variation. Unobserved time trends are not a concern because the control group showed no significant trend.

37 In addition, as a sensitivity, I re-estimated the tree excluding the awareness variable.
Beyond awareness, the treatment effect is estimated to be larger for households with higher baseline consumption, both in percentages and in levels. Among aware households, the estimated effect was -11.2% (node [3]) for households with baseline consumption of over 0.12 kWh per half-hour, compared to a noisy estimate of +8.4% (node [10]) for those below. This threshold of 0.12 kWh is very low, at the 5\textsuperscript{th} percentile of average baseline peak consumption. Baseline consumption also matters for unaware households, with households above the 0.25 kWh threshold (the 24\textsuperscript{th} percentile) reducing by -4.2% (node [12]) compared to -0.2% (node [13], not statistically significant) for those below it.\footnote{One concern is that such differences are spurious, in that larger users would naturally exhibit larger changes in levels. However, all treatment effects here in are percentage changes, implying that the effect is not limited to changes in levels.}

Among aware households with more than 0.12 kWh baseline consumption, the remaining sources of heterogeneity stem entirely from the information treatments. The in-home electricity display (IHD) produced the strongest effect on peak consumption (-14.6%, node [4]), followed by the monthly bill (-11.7%, node [6]), then by the bi-monthly bill with the overall load reduction (OLR) incentive (-10.7%, node [8]), then by the bi-monthly bill treatment (-7.9%, node [9]).

These differences in effects are likely due to the information treatments amplifying the effectiveness of time-of-use pricing, rather than having independent effects on consumption. This is because there is little effect on weekends and holidays, when time-of-use pricing was not applied but the information treatment remained.\footnote{See appendix for the treatment effects during weekends and holidays, when time-of-use pricing was not in effect.}

No other features of the data proved to be robustly related to treatment effect resulting tree splits on qualitatively similar variables, including the information treatment indicators and baseline energy consumption. The key distinction is that instead of splitting on the (now-excluded) “awareness” variable, the tree splits on the variable that is the single strongest predictor of awareness: access to internet in the home. In other words, the restricted tree is essentially approximating the same subgroups, but it is doing so with less precision because of the difficulty of using observables to predict awareness, which is documented in section 2.5.4.

\footnote{38} One concern is that such differences are spurious, in that larger users would naturally exhibit larger changes in levels. However, all treatment effects here in are percentage changes, implying that the effect is not limited to changes in levels.

\footnote{39} See appendix for the treatment effects during weekends and holidays, when time-of-use pricing was not in effect.

37
heterogeneity. This suggests that, conditional on these factors (awareness, baseline consumption, and information provision), no other observables are robustly related to heterogeneity in treatment effects.

I also estimate treatment effects during nighttime and off-peak daytime hours, during which prices were somewhat lower for the treatment group. The program increased nighttime consumption by +1.8% (with standard error of 1.1%) and decreased daytime consumption by -2.2% (with a standard error of 0.7%). The algorithm finds no heterogeneity in these treatment effects.

Implications

These results suggest that the key to designing effective time-of-use pricing policy lies in information and awareness, with some role for targeting households with non-trivial baseline consumption. Other than baseline consumption, the only dimensions that appear to matter are the awareness of the pricing policy and the amount of information provided to households. Conditional on these factors, no other source of heterogeneity is robust to cross-validation methods.

This suggests that complicated household targeting mechanisms are not necessary to produce good results. Conditional on awareness and information provision, baseline consumption appears to be a sufficient statistic for any other observables that may affect the impacts of time-of-use pricing.

In addition, the model finds no heterogeneity along tariff structure level, the other key dimension of treatment that the trial was designed to test. I turn to this in the next section.

40 While the first stage of the CT algorithm initially estimates a larger tree with richer heterogeneity than shown here, the second stage finds that those additional splits are not found to be robust to cross-validation. The tree shown in Figure 2.4 is optimally-sized tree according to 10-fold cross-validation.
Figure 2.4: Heterogeneous Treatment Effects on Consumption during Peak Periods

2.5.3 Treatment Effects by Tariff Levels

Estimated Demand Curves

The results above found no robust heterogeneity in effects other than awareness, information provision, and baseline consumption. This implies that the higher prices did not lead to significantly larger reductions, which may be surprising to economists. In this section, I explicitly investigate the effects of different tariff levels to confirm this previous finding.

In particular, I use the same difference-in-differences estimator from the previous section to estimate treatment effects by tariff group for peak, night, and daytime...
offpeak periods separately. I use these results to draw an average demand curve. This is possible because of the large variation in prices that households faced during peak periods. The results are shown in Figure 2.5.

The demand curve during peak periods, shown on the right-hand side of the figure, is very inelastic at high prices. The smallest price increase (from the control of 14.1 to 20 € cents per kWh) features a significant reduction, but further increases yield strongly diminishing returns. While the lowest tariff treatment (20 € cents) resulted in a statistically smaller effect than the higher tariff treatments did \(p = 0.013\), an F-test fails to reject the null that all four tariffs have jointly equal effects on peak consumption \(p = 0.08\). This is particularly surprising because, under the commonly-used assumption of a constant elasticity, reductions should scale proportionally with the magnitude of the price increase, but I find no evidence that it increases at all.

This rejection of a constant elasticity is not simply a result of weak statistical power, as the size of the standard errors rule out proportional increases. For example, the second peak tier (€26 cents, and increase of €11.9 cents) represents approximately twice as large of a price increase of the first tier (€20 cents, an increase of €5.9 cents). Given a constant elasticity, the second tier would be expected to lead to twice as large a reduction. An effect of this magnitude would be easily detectable at the 5% significance level given the estimated standard errors. In particular, any difference between these two tiers in excess of 40% would be statistically detectable at the 5% level, and that would nonetheless still be consistent with a declining elasticity.

The demand curve during daytime (off-peak) periods is in the middle of the fig-

\[\text{41 This result was already found by Commission for Energy Regulation (2011b).}\]

\[\text{42 In this data, the point estimate for the second tier is only 25\% larger than the that of the first tier, well within the confidence interval.}\]

40
There is little price variation during off-peak periods, but the effects during this period have a counterintuitive sign, with consumption decreasing despite lower prices. This was previously discussed in section 2.4 and is likely due to households preemptively turning off devices in anticipation of higher prices during peak periods. Here, it manifests as an apparently upward-sloping demand curve to contemporaneous prices.

During the night, there is a strongly inelastic demand curve, which can be seen in the lower left-hand side of the figure.

Altogether, it appears that the effects of time-of-use pricing depend on its mere existence, and not the level of the prices. This confirms the results from the previous section, which found that the causal tree algorithm found no apparent heterogeneity in the treatment effects by price level. There are several interpretations for this result. One interpretation is that households face strongly convex costs of reducing peak energy use, with small reductions being easy to achieve (turning off lights) but larger ones being very costly (turning off major appliances). An alternative interpretation is that households respond to the knowledge of differential pricing, but are not attentive to the specific level of the price.

There are several important caveats to this result. First, while the price increases during peak hours were large in percentage terms (up to 170%), they may not be large enough to command a household’s attention. For example, the highest tariff was 38 cents, up from 14.1 cents. This difference is less than 24 cents per kWh. On average, households in the sample consume 1.75 kWh during the 2-hour peak window each day, implying that the portion of their bill associated with peak pricing would only be about 42 cents per day (24 times 1.75), even if they did not respond at all. This small amount of money may not be enough to motivate reductions, suggesting much steeper price increases may be required to induce larger reductions.
A second caveat is that these estimates derive from the behavior of Irish residential consumers, and other types of consumers may respond differently. In particular, the households did not receive automation technology, such as programmable thermostats, that allows households to set price responsive behavior in advance. In addition, Irish electricity demand does not primarily derive from easily-automated home heating and cooling demand. Newer technology allows users set their thermostats to automatically adjust their set points in response to real-time prices (see, e.g. Bollinger and Hartmann (2016)). Households with access to this technology may exhibit more sensitivity to price levels. Therefore, these results suggest that the price level is not a strong driver of peak demand in the absence of automation technology. This suggests that automation technology may be necessary for real-time pricing to be more effective than time-of-use pricing.

**Implications of Declining Elasticities for Real-Time Pricing**

Regardless of the reason for this declining elasticity, it raises concern about the effectiveness of real-time pricing over TOU pricing in contexts similar to Irish households. If the exact price level does not matter—only that peak usage is priced differently—then real-time pricing is unlikely to change hour-by-hour electricity consumption. Hence, if real-time pricing does not affect behavior, then it will not reduce peak load when needed, raising concerns about whether it can achieve the goals it is designed.

To illustrate this point, I run a simulation of the effects of real-time pricing on consumption. I first estimate demand curves shown in Figure 2.5 separately for each “leaf” sub-group identified by the causal tree (Figure 2.4). These heterogeneous demand curves generally show similarly small and declining elasticities. Using these

---

43 Home heating in Ireland is typically fueled by oil or gas. Air conditioning is relatively uncommon in Ireland due to its climate. Water heating is typically powered by electricity, but on-demand heating is common in Ireland and is not as easily automated as home heating.

44 The exception is the group in node [10] which showed a statistically insignificant increase in
Notes: Each line represents estimated demand curves by period of the day (night, off-peak daytime, and peak). To construct these curves, I estimated treatment effects in percentage changes using a difference-in-differences (as in the causal tree estimates) for each tariff group and time period separately. I then transformed these percentage changes to average demand curves by multiplying the percentage change at each price level by the average baseline consumption in the relevant periods. Hence, these should be thought of as demand curves for the average consumer. 95% confidence intervals for the mean effects are shown, representing the uncertainty about the average treatment effects relative to the control group, which faced prices of 14.1 € cents.

**Figure 2.5:** Estimated Demand Curves, by Pricing Period

demand curves, I simulate individual household consumption under three different kinds of pricing: flat pricing (14.1 cents per kWh), TOU pricing as implemented in this experiment, and real-time wholesale spot electricity prices in Ireland.\textsuperscript{45}

consumption in response to higher prices, suggesting an upward-sloping demand curve. Because of the counterintuitive sign for this small subgroup, I treat this group as unresponsive for this simulation. I also treat daytime consumption as unresponsive, due to its similarly counterintuitive sign. Because the focus of this simulation is peak consumption, this choice is largely immaterial.

\textsuperscript{45} Spot prices represent “EP2” final prices collected from http://www.sem-o.com/marketdata/Pages/dynamicreports.aspx. I use the shadow prices (without capacity uplifts) because that represents true marginal generation costs, but including the uplift has very little effect on this simulation. Average spot prices are lower than average retail prices because they do not include transmission and distribution charges. I effectively assume that those costs would be recovered
Figure 2.6 shows total consumption across all simulated households for a day with particularly high demand and spot prices: January 4, 2010. Focusing on the peak periods, we see that TOU pricing generally reduces consumption relative to flat pricing (green versus blue lines) across the entire pricing window, including the hour of highest consumption. This indicates that TOU pricing successfully reduces peak load.

By contrast, real-time pricing does not reduce peak load and may even increase it. Wholesale electricity prices spiked that day, and the simulation suggests that this would have reduced consumption during the high-priced period. However, the declining elasticity implies that the response to this price spike would be very similar to the response to standard TOU pricing. Further, while real-time prices did indeed spike on this peak-load day, they did not spike during the peak hour. Large price spikes such as these are often fleeting because lower-cost suppliers cannot instantaneously ramp up to meet demand, but can do so on relatively short notice. In other words, because of adjustments on the supply side, the period of peak demand is not always the same as the period of peak prices. In this example, during the period of peak demand, the market price was lower than the flat retail price. As a result, households would be charged lower prices during this period under real time pricing than they would under a flat tariff, resulting in even higher peak load. This shows how real-time pricing can actually be strictly worse than flat-rate pricing from the perspective of peak load.

Of course, Figure 2.6 assumes that households would respond instantaneously to rapidly changing prices under a real-time pricing regime. This is clearly a strong assumption, but the alternative assumption of a sluggish response would make real-time pricing appear even less effective. Further, a more realistic assumption is that households respond less to real-time pricing than to predictable TOU pricing, since
TOU prices are set and known in advance. Real-time prices change every half hour in Ireland and so would require households to devote constant attention to their meters, which would be very costly and infeasible for many households. If households would not respond instantaneously, the impact of hour-to-hour changes in prices would be more muted than TOU prices would be, making real-time pricing even less effective at affecting behavior than the simulation suggests.

Figure 2.6: Simulated Consumption under Different Pricing Regimes: January 4, 2010

2.5.4 Who Is Likely to Be Aware?

Since awareness of time-of-use pricing is key to its effectiveness, is it possible to identify *ex ante* which households are likely to be aware of the policy? If so, policy can be targeted towards those households to improve outcomes. I answer this question
using three different methods, all of which reach the same conclusion: household observables have only small predictive power for awareness.

The first method is involves estimating a linear probability model of awareness as a function of pre-treatment household observables and treatment group indicators.\textsuperscript{46} The truncated set of the results is shown in column (1) Table 2.4 (all significant variables are shown, and the other 98 variables are suppressed on account of space). Of the 123 variables considered, 14 (11\%) are significant at the standard 5\% level. This is somewhat larger than the 5\% expected if there were no explanatory power of the observables, but the low share of significant coefficients suggests caution in over-interpreting the ones that do turn out significant. In addition, no coefficients are significant with a Bonferroni correction.

To determine how many of these significant coefficients are spurious due to sampling variability, I employ a machine learning method, post-selection Lasso regression, to eliminate spurious variables from the model. This involves estimating the linear probability model with a penalty term for large and non-zero coefficient estimates to determine which variables to retain in the model.\textsuperscript{47}

\textsuperscript{46} Only treatment households are included in this analysis, as they are the relevant group. A logistic model produces very similar results for this entire section, but its coefficients are more difficult for the reader to interpret. The typical disadvantage of linear probability models—that they may imply predicted probabilities outside of the interval \([0, 1]\) is not a significant problem in this case, as the vast majority of predicted probabilities fall within that interval.

\textsuperscript{47} Lasso regressions are estimated by minimizing a standard loss function (such as the sum of the sum of square residuals or negative log likelihood) plus a penalty term proportional to the L1-norm (\(|\beta|\)) of the standardized coefficient vector (here denoted \(\beta\)). In the context of a linear model, the objective function is

\[
\min_{\beta} \sum_{i} (y_i - X_i \beta)^2 + \lambda |\beta|\]

where all \(X_i\) variables have been standardized to have a mean of zero and a standard deviation of one. Depending on the value of \(\lambda\), the results range from collapsing to exactly the OLS estimates (when \(\lambda = 0\)) to a null model with no covariates (for sufficiently large \(\lambda\)). The optimal value of \(\lambda\) is chosen by estimating the out-of-sample prediction error through cross-validation. This process completely eliminates some variables of the model, while retaining ones that have out-of-sample predictive power.

I use post-selection Lasso, which means estimating a Lasso model, determining which variables are retained, then re-estimating the model by OLS using only those retained variables.
The magnitude of the penalty parameter is chosen through cross-validation. I present results using two standard criteria for the optimal penalty parameter. First, column (2) of Table 2.4 shows the results using the penalty parameter that minimizes the out-of-sample root mean square error (RMSE) according to cross validation. This method eliminates 86% of the variables from the model, finding that they are ineffective at predicting consumer awareness, and the ones that remain have small magnitudes.

The second standard method for choosing the Lasso penalty parameter is to use the most parsimonious model (i.e., the largest penalty parameter) that produces an out-of-sample RMSE that is within one standard error of the minimum RMSE. This penalty parameter results in the much simpler model shown in column (3). This model drops all variables except for an indicator for whether the household has internet access in the home. Those with internet access are 13 percentage points more likely to be aware of their pricing regime (91% versus 78%). While a 13 percentage point difference is non-trivial, it is somewhat modest given that it is the most important predictor according to the Lasso results. Altogether, this suggests that there is not a clearly reliable way to predict whether households will be aware of the TOU pricing ex ante.

To further illustrate this weak power that observables have to predict awareness, Figure 2.7 shows histograms of fitted probabilities using each model in Table 2.4. These are the probabilities one would want to compute if policymakers wanted to

---

48 This is motivated by the fact that the model produces a statistically indistinguishable error rate, but is more parsimonious and so avoids overfitting. See James et al. (2013).

49 Regression trees also find that internet access is the most important predictor for awareness, but cross validation excludes it as non-robust.

50 This figure truncates fitted probabilities from the linear model that are greater than 100%, replacing them with 100%. Results from a logistic regression, which do not feature fitted probabilities exceeding 1, look similar. In addition, histograms of the fitted probabilities from a much more flexible random forest model are also similar, implying that these results are not the result of an insufficiently flexible functional form.
target TOU pricing on observables, perhaps by preferentially treating those who are more likely to be aware. In general, they show that few households exhibit low probabilities of awareness, suggesting that ignoring such households in a policy rollout would have little effect. In the full model, the lowest predicted probability of awareness is about 40%, and only 0.4% of households are more likely to be unaware than aware.\textsuperscript{51} In the RMSE-minimizing Lasso model, the lowest predicted probability of awareness is about 60%, implying that only targeting households that are more likely to be aware than unaware would have no effect on the treatment. The simplest model (the Lasso using the 1 standard error rule for the penalty parameter) leads to only two predicted probabilities: 91% for those with internet access, and 78% for those without. All of these results suggest little role for using observables to target TOU pricing towards households likely to be aware.

\section{Conclusion}

I apply and extend new machine learning methods to estimate heterogeneous treatment effects from time-of-use pricing and information provision on residential electricity consumption. Most importantly, the effect of time-of-use pricing on peak energy consumption is 4.5 times larger for households who are aware of the change in their pricing structure compared to those who are not (-10\% versus -2.3\%). Beyond awareness, the treatment effect varies significantly based on baseline electricity consumption and the amount of information provided to consumers about their dynamic pricing and energy usage. Households with very low baseline energy use do not reduce their consumption on average. Meanwhile, the strongest information treatment (an in-home electricity monitor) leads to nearly twice the reduction in peak consumption (-15\%) compared to the weakest information treatment (a bi-monthly

\textsuperscript{51} That is, in the full model, only 0.4\% of households have a probability of awareness under 50\% according to their observables.
energy use statement, -8%).

According to cross-validation techniques, no other observables—or permutations thereof—are robustly related to treatment effect heterogeneity (conditional on the above factors of awareness, baseline consumption, and information treatment). This includes considering potential heterogeneity on more than 150 observables encompassing demographics, house attributes, household appliance ownership, and attitudes towards energy and environmental issues. This suggests that baseline consumption acts as a “sufficient statistic”, and obtaining further detailed household characteristics is unnecessary for predicting differential responses.

Larger price increases do not induce significantly larger responses, which suggests that fine-tuning retail pricing is unlikely to result in incremental demand response.

Awareness is not reliably predictable even with rich information about observable household characteristics, which suggests that attempting to target TOU pricing based on awareness is unlikely to prove fruitful, although targeting on baseline consumption could be effective.

Altogether, these results suggest that the significant attention economists devote to “getting the prices right” is less important than getting consumers to pay attention in the first place.
Table 2.4: Post-Selection Lasso Linear Probability Model of Awareness

<table>
<thead>
<tr>
<th>Dependent variable: Aware of Tariff Change (Indicator)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet Access in Home (Indicator)</td>
<td>0.06</td>
<td>0.06</td>
<td>0.13</td>
</tr>
<tr>
<td>Use Internet Regularly (Indicator)</td>
<td>0.02</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Number of Children under 15 in Home</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Already Made Changes to Reduce Electricity Use (Pre-Trial)</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Home Heat: Oil (Indicator; “none” omitted)</td>
<td>−0.06</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Water Heat: Oil (Indicator; “none” omitted)</td>
<td>0.05</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Number of Dishwashers</td>
<td>0.03</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Number of Desktop Computers</td>
<td>0.02</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Share of Windows Double Glazed (0-5 = 0/25/50/75/100%)</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Expect Participating in Trial Will Reduce My Bill (Indicator)</td>
<td>0.05</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Expect to Choose More Efficient Appliances*</td>
<td>−0.03</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Satisfied with Share of Electricity From Renewables*</td>
<td>0.04</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Female Respondent (Indicator)</td>
<td>0.04</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Social Class: AB, Manager/Professional (Indicator; “Refused” omitted)</td>
<td>−0.02</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Education: Third (e.g., University) (Indicator; “Refused” omitted)</td>
<td>0.05</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Own home with mortgage (Indicator; “Rent from private owner” omitted)</td>
<td>0.06</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Cook stove type: Electric (Indicator; solid fuel omitted)</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cook stove type: Gas (Indicator; solid fuel omitted)</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Info. Treatment: Monthly bill (Indicator; Bi-monthly omitted)</td>
<td>0.04</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Info. Treatment: In-Home Display (Indicator; Bi-monthly omitted)</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Info. Treatment: OLI (Indicator; Bi-monthly omitted)</td>
<td>−0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tariff Treatment: B (Indicator; A omitted)</td>
<td>−0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tariff Treatment: C (Indicator; A omitted)</td>
<td>−0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tariff Treatment: D (Indicator; A omitted)</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.96</td>
<td>0.34</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Observations 2,328 2,328 2,328
R² 0.12 0.07 0.03
Adjusted R² 0.07 0.07 0.03
F Statistic 24.4 10.53 74.23

Number of Covariates 123 17 1
Number of Covariates Not Shown 98 na na
Number of Covariates Significant (5% level) 14 na na
Share of Covariates Significant (5% level) 11.4% na na

Notes: Due to space limitations, only a subset of the 123 included variables in column (1) are shown. The variables shown include all statistically significant ones at the 5% level, all variables retained by the Lasso technique, and all treatment group indicators. The full set of results is available upon request. Column (2) shows the post-Lasso selected model using the model that minimizes the out-of-sample root mean square error (RMSE). Column (3) shows the most parsimonious model that has an error within one standard error of the RMSE of the RMSE-minimizing model shown in (2). p-values and significance levels not presented for columns (2) and (3) because classical tests are not applicable the post-selection model. Methods for inference in post-selection models is a contrarian area of on-going research. Baseline Average Consumption variables units are kWh per 30 minute interval. Approximately 78 percent of all households have internet in their home. In column (1), the excluded indicators for social class and education are “refused to answer”, and the excluded indicator for “own home with mortgage” is “rent from private owner”. In columns (2) and (3), the excluded indicators are all other possible outcomes.

* These variables featured numeric responses, where respondents reported on a 1-5 scale to what extent they agree (1) or disagree (5) with the statement.

This table was generated using the stargazer package (Hlavac 2015) for R.
Figure 2.7: Histograms of Fitted Probabilities of Awareness, from Models shown in Table 2.4
Chapter 3

Prices versus Quantities with Policy Updating
(Co-author: William A. Pizer)

Note: This work was co-authored with William A. Pizer. Prest’s contribution involved conception of the initial idea, development of the basic modeling framework, derivation of the initial results, and authoring a first draft. In addition, modeling of political “noise”, development of the $T$-period model, and writing and editing subsequent drafts were undertaken collaboratively.

3.1 Introduction

For more than forty years, economic thinking about the relative welfare advantage of alternative price and quantity regulation under uncertainty has gone something like this: When the regulator is uncertain about costs known to firms, the difference in expected net benefits between otherwise equivalent price and quantity policies hinges on the difference between the slopes of the marginal benefit and cost schedules, multiplied by the variance of the cost shock Weitzman (1974). One intuition is that government policy is attempting to replicate society’s expected marginal benefits in the form of a demand schedule in the regulated market. A flat schedule (prices)
works better when marginal benefits are relatively flat (e.g., Pizer 2002); a vertical schedule (quantities) works better when marginal benefits are relatively steep. A corollary to the Weitzman result is that uncertainty about marginal benefits (unless correlated with costs) does not affect comparative advantage of the two instruments (a point made in a footnote in the original paper, but elaborated upon in Stavins 1996).

As we show in this paper, however, the outcomes can change significantly in a dynamic setting where policy changes over time. In a model that extends Weitzman to multiple periods with policy updating, we show how intertemporally tradable quantity regulation can achieve the first-best outcome. Price regulation can only achieve the first best in special cases and is never strictly preferred in this setting. Moreover, the comparative advantage now depends on uncertainty about both costs and benefits. We obtain this result despite the maintained assumption that the government sets its policy for each period based on previous-period information, while the firm chooses its action based on current-period information.

This result arises because of the interaction between policy updating and the arbitrage condition created by an intertemporally tradable quantity regulation. In the static model, cost shocks lead price and quantity regulation to diverge in different ways from the first best. The same outcomes would occur in our dynamic setting if the quantity instrument could not be traded across periods. With intertemporal trading, however, the opportunity to save (or borrow) regulated quantities for (or from) the next period implies an arbitrage condition between current and expected next-period prices as firms cost minimize. A benefit-maximizing government can take advantage of this behavior with a particular strategy of updating the quantity regulation in order to achieve the first best. It can do this even with its constraint of setting policy each period based on previous-period information.

By contrast, under price regulation there is no comparable arbitrage condition.
The government’s strategy of updating the price regulation can influence expectations about next-period prices and policies, but cost-minimizing behavior by firms does not change the current-period price that they face. That is, even if firms shift polluting activity in response to expectations about future pollution taxes, they do not affect the current pollution tax. Coupled with the assumption that the current-period policy is set based on previous-period information, the best that a benefit-maximizing government can do with a price policy in this setting is to achieve a second-best outcome similar to the result in Weitzman (1974). A price regulator can only attain the benefits of intertemporally tradable quantity regulation by setting prices retroactively, something not typically seen in practice. Indeed, a retroactive tax arguably amounts to an *ex post facto* law that is prohibited in many countries.

Viewed another way, a dynamic policy setting can relax the information requirements needed to achieve the first-best policy. An intertemporally tradable quantity policy need not be set to achieve the first-best outcome in any particular period so long as firms expect the quantity allocation to be adjusted appropriately in the future. In a simple two-period model without discounting, we show how the government merely sets the first-best allocation in the second period, plus a “top-up” for the first period, and achieves the first-best outcome in both periods. This parallels actual experience with tradable permit programs (e.g., California and Europe), which often provide generous permit allocations in initial phases of the program, followed by later adjustments to the caps.

Of course it far from obvious whether policy updates to a particular regulation are always driven by society’s welfare maximization. With this in mind, our paper considers the possibility that the regulator’s objective deviates from society’s marginal net benefits by some amount. This could reflect special interest politics, as environmental advocates or business interests hold too much sway over the government, or legal or other constraints. This possibility — what we might call uncertain
political “noise” — favors price regulation as the adverse influence is at least delayed until the policy takes effect.

The motivation for this line of thinking stems from an attempt to explain observed outcomes under real-world quantity regulation and to relate these outcomes to the question of comparative advantage. Figure 3.1 shows the price behavior in the SO$_2$ trading program from 1995 through 2010. Prior to 2004, we might have told a traditional Weitzman story. The original prediction was a steady-state price of around $600 per ton for an emission level of 9 million tons, but various favorable cost shocks led to an equilibrium around $200. Thus, a Weitzman-style deadweight loss could be computed based on fixing quantities rather than prices at the \textit{ex ante}
optimum and having marginal costs end up $400 lower.\(^1\) Such a calculation would involve comparing the slopes of marginal costs and benefits, but would likely support price-based regulation due to relatively flat marginal benefits.\(^2\)

Starting in 2004, a different picture emerges. Future reforms to the program — a policy update in the form of the “Clean Air Interstate Rule” or CAIR — began to be debated. This new rule would update the existing policy by halving the emission limit to about 4 million tons. The update was primarily driven by new evidence of large health benefits from lower emissions. However, it also considered the low compliance costs observed with the original cap.

The EPA projected a steady-state price of $1,400 under the new rule, with 2010 prices projected to be $700.\(^3\) None of the changes to the program would have come into effect prior to 2010. Yet, the current (spot) price of SO\(_2\) emission permits arguably moved in parallel to shifting expectations of future prices because of the ability to save or “bank” current permits for future use. In turn, this created incentives for additional emission reductions and cleaner air that were in line with the (improved) marginal benefit estimate, achieving higher societal net benefits in 2004. Had an SO\(_2\) tax been in place, however, an expectation of a higher tax in 2010 would not have influenced the current tax. This feature arguably favors quantity regulation, as the gains from planned improvements in the policy target are realized immediately, before the change has actually been implemented.

Starting in 2006, however, the price fell as the EPA reconsidered the proposed rules and later court challenges succeeded in vacating the regulation. In this case, particularly as the price fell to zero well before the new policy came into effect, an

\(^1\) See Schmalensee et al. (1998) for broader discussion of SO\(_2\) trading program and early expectations concerning allowance prices.

\(^2\) Pope et al. (2015) suggest marginal benefits from pollution reduction may not decline at lower pollution levels and may even increase.

\(^3\) See Environmental Protection Agency (2005), Table 7-3.
SO₂ tax would have at least maintained the previous level of net benefits until any policy update occurred.⁴ Here, price regulation arguably would be preferred.

In contrast to the pre-2004 period in this market, the distinction between price and quantity regulation over 2004-2010 is related to the equalization of current and expected future prices under quantity regulation. The comparative advantage of prices versus quantities would then depend on the degree to which future policy updates and prices reflect welfare maximization or something else. More generally, it also matters that intertemporal trading and arbitrage occurs and that market behavior is dominated by attention to such updates.

Among existing market-based environmental regulations, the SO₂ program appears far from unique in terms of intertemporal trading and policy updates. The linkage over time of markets for regulated quantities through permit banking is consistent with virtually all existing tradable permit programs.⁵ Meanwhile, most real-world environmental regulations are updated over time in response to new information. The 1970 Clean Air Act makes specific provisions for the National Ambient Air Quality Standards to be updated every five years in response to new information (see 42 U.S. Code §7409). More recently, the European Union announced a package of changes in their Emissions Trading Scheme, in large part owing to low prices and (presumably) low compliance costs.⁶ California is similarly contemplating revised targets for the next phases of their program, hopefully based on improved informa-

⁴ It is an open question whether marginal cost really declined to zero — e.g., no effort being made to reduce emissions — or whether emissions were being reduced based on other regulatory constraints and there was simply an excess supply of permits over the remaining horizon of the (now defunct) policy. In any case, it seems fair to conclude that there was less effort occurring when the price was zero than when it was $200 per ton.

⁵ This includes lead phase-down program in the 1980s, the acid rain program in the 1990s and 2000s, the European Emissions Trading Scheme, California’s cap-and-trade program, among others. Most programs allow limited borrowing. In practice, most programs feature significant positive bank balances, suggesting that borrowing limitations have generally been non-binding.

tion. At the U.S. federal level, in 2010 an interagency working group estimated the social cost of carbon (SCC)\(^8\) to be $24 per metric ton in 2010. That number was revised upward to $37 based on new modeling estimates of damages in 2013.\(^9\) Finally, we note that the Paris Agreement, negotiated in 2015 to address global climate change, requires countries to submit new policy proposals every five years.\(^10\)

There is a similar list of policy updates where one might question whether aggregate welfare maximization is the main goal. In 2017, the Trump administration proposed revising the SCC down to as low as $1 per ton, a level well below the range of consensus estimates.\(^11\) Other examples include New Jersey’s decision to exit the Regional Greenhouse Gas Initiative (RGGI) in May 2011, Australia’s decision to terminate their carbon pricing program in July 2014, and the U.S. Supreme Court decision to stay implementation of the Clean Power Plan in February 2016.

In the remainder of the paper, we first review the basic Weitzman (1974) result and other relevant literature. We then show how the Weitzman result changes if we allow policy updates and intertemporal quantity trading. To illustrate our main results, we present a simple two-period model with a common shock across both periods and no discounting.

After showing how intertemporally tradable quantity regulation can achieve the first best outcome, we introduce the possibility that policy updates might be driven by influences other than new information about costs and benefits, such as shifting special interests or evolving legal or other constraints. This reintroduces the pos-

---

\(^7\) See [http://www.arb.ca.gov/cc/scopingplan/scopingplan.htm](http://www.arb.ca.gov/cc/scopingplan/scopingplan.htm).

\(^8\) The SCC represents the marginal benefit of reducing CO\(_2\) emissions by one metric ton.


\(^10\) See Article 4, paragraph 9.

sibility that price regulation is strictly preferred, with the comparative advantage determined by the magnitude of this policy noise relative to new information about costs and benefits. All of these results can be generalized to multiple periods with AR(1) error processes and discounting, as we demonstrate in section 3.5.

Finally, we apply our model to the case of global climate change, estimating that intertemporally tradable quantities provide an expected welfare gain over price regulation of $2 billion per five-year period, assuming policies are updated every five years, a number that grows if updates occur less frequently.

3.2 The Weitzman Result and Subsequent Work

In his seminal 1974 article, Martin Weitzman explored the difference between price and quantity regulation of some outcome $q$ (a good or a bad)\(^{12}\) in a simple linear model. Assuming costs and benefits represented by

\[
C(q, \theta) = c_0 + (c_1 + \theta)(q - \hat{q}) + \frac{c_2}{2}(q - \hat{q})^2, \tag{3.1}
\]

\[
B(q, \eta) = b_0 + (b_1 + \eta)(q - \hat{q}) - \frac{b_2}{2}(q - \hat{q})^2, \tag{3.2}
\]

linear marginal costs and benefits can be written as

\[
MC(q, \theta) = c_1 + \theta + c_2(q - \hat{q}) \tag{3.3}
\]

\[
MB(q, \eta) = b_1 + \eta - b_2(q - \hat{q}). \tag{3.4}
\]

with $E[\theta] = E[\eta] = 0$. Here, $\theta$ represents cost uncertainty while benefit uncertainty is captured in a similar way by $\eta$. For an arbitrary $\hat{q}$, the parameters $c_1$ and $c_2$ describe marginal costs in a flexible way, and similarly so for $b_1$ and $b_2$. As in

\(^{12}\) Weitzman focuses on $q$ as a good (e.g., clean air) for simplicity. However, our motivating example is pollution where marginal benefits may be flat and where quantity controls are more easily viewed as regulating a bad (e.g., emissions). There is nothing in the model or notation that favors one or the other, but the reader should keep in mind that for bads, marginal benefits and hence prices are negative; firms must pay rather than being paid for producing bads.
Weitzman, we assume $c_2 > 0$ and $b_2 \geq 0$, and $MC(0, \theta) < MB(0, \eta)$ for all $\theta$ and $\eta$ and $MC(q, \theta) > MB(q, \eta)$ for large enough $q$. This implies a single first-best outcome, regardless of whether $q$ is a good or a bad.

In his paper, the linear model is described as an approximation to more general functions about the point $\hat{q}$ where marginal costs equal marginal benefits on average. That is, he (and we) assume $MC(\hat{q}, 0) = MB(\hat{q}, 0)$, or equivalently $c_1 = b_1$. While this assumption is not required to derive the key results, it simplifies the exposition and the expected optimal government policies. These become $\hat{q}^O = \hat{q}$ for quantities and $\hat{p}^O = c_1 = b_1$ for prices. Here and throughout, we use a tilde over a variable to indicate a chosen government policy; the superscript $O$ more specifically refers to the Original Weitzman framework for policymaking.

With this setup, Weitzman derives his main result. When price and quantity regulations are set to maximize expected net benefits, the expected welfare advantage of prices is given by

$$\Delta^O = \frac{\sigma^2_\theta}{2c_2^2}(c_2 - b_2).$$

(3.5)

Weitzman’s appealingly simple result (3.5) shows that the direction of the welfare advantage $\Delta^O$ depends only on the relative slopes of marginal costs and marginal benefits, $c_2$ and $b_2$. We previously mentioned the intuition that a price policy, filling in the otherwise absent market demand role, better matches the marginal benefit schedule when it is relatively flat (and vice-versa for quantity policies when the marginal benefit schedule is relatively steep). Another intuition is that price regulation offers welfare savings because such a policy does more when costs are low and less when costs are high. To the extent costs are convex (large $c_2$), this improves welfare. At the same time, a price policy introduces variability into $q$, which is fixed.

---

13 This assumption is required for an interior solution where the optimal $q > 0$. This implies $\theta - \eta < (b_1 + b_2\hat{q}) - (c_1 - c_2\hat{q})$. That is, uncertainty is relatively small compared to the difference in marginal costs and benefits when $q = 0$. 
under quantity regulation. To the extent benefits are concave (large $b_2$), variable $q$ reduces welfare. The net effect depends on the sign of $c_2 - b_2$. Put simply, if marginal benefits are relatively steep compared to marginal costs, then getting the quantity “right” is important in welfare terms, implying that quantity instruments are expected to perform better. Otherwise, price instruments are to be preferred.

The magnitude of both the expected cost saving and benefit loss depends on variability in costs — here, the variation in $\theta$. Benefit uncertainty — variability in $\eta$ — does not appear in the expression. Benefit uncertainty does lower expected welfare, but it does so equally for price and quantity regulation.

Weitzman’s findings have been extended to many other contexts in the literature. This includes alternative forms of uncertainty (Fishelson 1976; Yohe 1978 and Stranlund and Ben-Haim 2008), market power (Moledina et al. 2003), and non-linear marginal costs and benefits (Yohe 1978; Kelly 2005). In all of these cases, the basic result generally remains that flatter marginal benefits favor prices and steeper marginal benefits favor quantities. Other work has considered the correlation between marginal cost and benefit uncertainty (Yohe 1978; Stavins 1996). Such correlation can overwhelm the basic Weitzman intuition in theory, but it has never been demonstrated to be relevant in practice. Kaplow and Shavell (2002) take a different tack. They propose to fix the price only after firms report their total emissions based on the expected marginal benefit schedule. This allows them to achieve welfare improvements over either fixed prices or fixed quantities, but raises questions as such policies have not been seen in practice. Kollenberg and Taschini (2016) consider how to best design a quantity mechanism to adjust over time depending on the realized size of the bank, effectively allowing a quantity policy to adjust based on realized shocks. Yet other work has examined whether the choice of price or quantity controls affect other outcomes (e.g., investment in Chao and Wilson 1993 and Zhao 2003), but none of this latter work speaks to the normative question of welfare impacts.
Most relevant for our paper are extensions to Weitzman that look at uncertainty and instrument choice in a dynamic policy context. Yates and Cronshaw (2001) consider the value of intertemporal permit trading in a deterministic setting, noting that allowing one-for-one banking may not be optimal and introduce the idea of trading ratios (something we return to in our multi-period model). Newell and Pizer (2003) and Fell et al. (2012) both extend the original Weitzman framework to multiple periods where benefits can depend on the accumulated level of the pollutant, rather than the annual flow. Despite the key difference that one allows intertemporal trading and the other does not, they find results broadly similar to Weitzman. Flatter marginal benefits and steeper marginal costs still favor prices, and vice-versa for quantity regulation. Importantly, these papers do not consider the possibility that policies might be updated.

Several papers do consider policy updates in a multi-period setting. Heutel (2012) considers how policies should be updated in response to the business cycle, but the model does not feature asymmetric information. Boleslavsky and Kelly (2014) consider how a government can adjust the stringency of policy over time to induce the firm to reveal its costs. Newell et al. (2005) are entirely focused on policy updating, but only in the limited sense of whether an intertemporally tradable quantity policy could respond to cost shocks in order to mimic a price policy over time. The idea of benefit uncertainty and that policy updating might be focused on new information about benefits is absent. In an unrelated paper focused on the idea of an allowance reserve (a hybrid price-quantity policy), Murray et al. (2009) argue that emissions trading with an allowance reserve could provide a higher level of cost-effectiveness than either emission taxes or emissions trading alone. Part of their argument is related to our idea that intertemporal emissions trading has an advantage when policies are updated in response to meaningful new information. Like the previous paper, this abstracts from the notion of welfare and benefits. They assume a cumulative
emission goal that is updated in the future and only seek to minimize costs.

Finally, Hoel and Karp (2002) compare welfare outcomes under price and quantity regulation with policy updates but without intertemporal trading of the quantity instrument. They focus on a stock pollutant, whereby the fluctuations in the accumulated stock that occur under a price policy can be attenuated with policy updates. This feature coupled with the lack of intertemporal trading of quantities gives prices an extra welfare advantage in this setting.

These papers focused on a multi-period setting can be categorized by whether they include four features: uncertainty, intertemporal trading, policy updating, and explicit welfare analysis. None of them contain all four, which leads us to the current paper. This paper can be viewed as extending and clarifying the notion that policy updates coupled with intertemporal trading seems to provide some sort of advantage for quantity controls. In contrast to previous work that included both of these features, however, we consider benefits and benefit uncertainty and focus on welfare. This allows us to show that policy updates seeking to maximize societal benefits lead to a clear preference for quantity regulation that can be traded over time.

### 3.3 Policy Updating

We now expand the earlier Weitzman model to two periods with the same, but uncertain, costs and benefits in both periods. The regulator sets the first-period price or quantity policy prior to learning about costs and benefit uncertainty, $\theta$ and $\eta$ in (3.1) and (3.2), but we consider the possibility that s/he sets the second-period policy once the uncertainty is known. That is, we change the information asymmetry to represent a lag in policy implementation in the context of evolving knowledge rather than a perpetual difference in what may be known or unknown.

We pause for a moment to highlight this slight shift in interpretation. Weitzman
focuses on the difficulty of the government to know true costs that are more naturally known to the firms that experience them. In the same way, one could view benefit uncertainty as a deep uncertainty about the true marginal benefits that the government simply cannot know. From such a vantage point, the government only can infer $\theta$ after firms respond to the policy and may never learn $\eta$ at all.

We take a slightly different perspective. We think of $\theta$ and $\eta$ as an evolution of society’s best estimate of marginal costs and benefits over time. That is, everyone’s best guess is $\theta = \eta = 0$ in period zero. Information improves in period one. This may be due to firms’ responses to the government policy but is also a function of improved science and experience that arises from the steady march of time. In our view the government learns about $\theta$ and $\eta$ alongside the firm but is unable or unwilling to act quickly enough to change the first-period policy directly, perhaps due to a relatively slow-moving legislative or bureaucratic process. This is an important assumption: If the government is able to act rapidly as information is known, there is no information asymmetry. However, we feel the assumption is justified based on observation of actual government policymaking (e.g., the five-year lag between the proposal of CAIR and its planned implementation date, during which the market reacted strongly).\(^{14}\)

With this information model, the government is able to set second-period policy based on this observed information on costs and benefits ($\theta$ and $\eta$) and does so according to a pre-announced rule. Meanwhile, the firm can deduce the government’s second-period policy before taking action in the first period through knowledge of $\theta$ and $\eta$ and this pre-announced rule.\(^{15}\) The information structure and decisions over

\(^{14}\) We intentionally do not specify the length of each period. As in Hoel and Karp (2002), the length of each period primarily determines how often policies are updated. To the extent policymaking is a slow process, it is perhaps best to imagine a period being several or even many years long. In this case, it makes sense to assume first-period policy will not be changed directly.

\(^{15}\) In this case, we are assuming the government is a Stackelberg leader and the firm does not attempt to strategically influence the government’s second-period behavior. For example, we might
time are depicted in Figure 3.2.

![Timeline of events with updating in a simple two-period model](image)

**Figure 3.2:** Timeline of events with updating in a simple two-period model

In addition to the assumed information structure and decision sequence, we make the further assumption that the quantity instrument can be traded between the two periods. As we have noted, this is consistent with virtually all observed tradable permit systems that allow current period permits to be saved, or “banked”, for use in the future and many that allow some volume of future period permits to be used early, or “borrowed”. While many programs feature some limitations on borrowing, in practice they have generally experienced significant positive bank balances. While our model assumes no borrowing constraints, such limitations have typically been non-binding in practice.

Also consistent with observed policy designs, we assume the regulator sets policy (either quantity or price) for the future and not retroactively. If we allow the imagine a large firm trying to manipulate the government’s view about the cost shock (which otherwise can be determined by the market price and the observed $q$). This is consistent with the idea of the regulated “firm” representing an aggregation of multiple firms in a competitive market.

---

16 By their very nature, price instruments that specify a tax or subsidy rate to be applied over a particular period of time cannot be traded between periods. Some have suggested mechanisms to imbue price mechanisms with intertemporal features. This includes setting price instruments retroactively or allowing firms the choice of paying either the current or future tax rate on current emissions. The first option, retroactive taxation, is not seen in practice over long time periods. The latter option would only be effective when the policy update sought to lower prices. Firms would not voluntarily pay a higher price.
regulator to set policies retroactively, the problem becomes trivial and the regulator simply waits for all uncertainty to be resolved before setting any policy. In reality, policymakers do not have this option. Indeed, retroactive policies are not typically seen in practice. While there are examples of tax changes enacted late in a given year being applied retroactively to liabilities incurred earlier in that year, using a retroactive mechanism over any substantial period of time seems infeasible in practice. Further, there is a question about the legality of retroactive policies in many legal systems, since they arguably amount to *ex post facto* laws.

We now characterize firm behavior, the government’s welfare-maximizing price and quantity policies, and the relative advantage of prices versus quantities, for both fixed and updated policies.

### 3.3.1 Firm Behavior

In the traditional Weitzman framework, firm behavior under a generic quantity policy $\tilde{q}$ was to trivially set $q = \tilde{q}$. Here, it is not so simple in the first period as the firm has the option to use the assigned quantities $\tilde{q}_1$ or to deviate, creating a bank $B = \tilde{q}_1 - q_1$. This bank can be in surplus or deficit $B \geq 0$ to be made up in the second period. For that reason, we work backwards: In the second period, there is no option except $q_2 = \tilde{q}_2 + B$. Then, in the first period we can write the firm’s problem as:

$$\max_B -C(\tilde{q}_1 - B, \theta) - C(\tilde{q}_2 + B, \theta).$$

We ignore discounting for the moment. We also take advantage of the fact that there is no uncertainty from the perspective of the firm. It knows the cost and benefit shocks before it makes any decisions and can deduce the government’s second-period policy based on $\theta$ and $\eta$. This leads to an arbitrage condition for the firm: $MC(q_1, \theta) = MC(q_2, \theta)$. That is, equalize marginal costs across periods. Given the
same cost function in both periods, we have a simple solution:

\[ q_1 = q_2 = \frac{\tilde{q}_1 + \tilde{q}_2}{2}. \]  

(3.6)

That is, divide the cumulative quantity regulation equally over the two periods. This is true regardless of what kind of quantity policy is used to define \( \tilde{q}_2 \).

Firm behavior under a price policy is the same as derived by Weitzman. Facing prices \( \tilde{p}_t \) in either period for output \( q_t \), firms will choose \( q_t \) to maximize profits. That is,

\[ \max_{q_t} \tilde{p}_t q_t - C(q_t, \theta), \]

recalling that \( \tilde{p}_t \) will be positive or negative depending on whether \( q_t \) is a good or a bad. This leads to a simple rule for the firm: \( MC(q_t, \theta) = \tilde{p}_t \). That is, set marginal costs each period to the regulated price. Given the definition of marginal costs in (3.3), this yields Weitzman’s response function,

\[ h(\tilde{p}_t, \theta) = \hat{q} + \frac{\tilde{p}_t - c_1 - \theta}{c_2}. \]  

(3.7)

### 3.3.2 Optimal Price and Quantity Policies

For comparison, we first consider what happens when the government chooses its policies to maximize expected net benefits over both periods without knowing the values of \( \eta \) and \( \theta \) (that is, without updating). Absent updating, this replicates the earlier Weitzman policy result but now extended over two periods, which we denote with superscript \( W \). In the case of intertemporally tradable quantities, it is clear from equation (3.6) that only the total quantity volume \( \tilde{q} = \tilde{q}_1 + \tilde{q}_2 \) matters as a policy choice variable and the allocation between periods does not matter. Following the Weitzman approach that costs and benefits are approximated around \( \hat{q} \) where \( b_1 = c_1 \), it is trivial to show that expected net benefits are maximized when the
cumulative quantity is given by \( \tilde{q}^W = 2\hat{q} \), which could be implemented with \( \tilde{q}_t^W = \hat{q} \) in each period. In the case of prices, the symmetry of the problem leads to the original Weitzman solution \( \tilde{p}_t^W = c_1 = b_1 \), resulting in equilibrium quantities of \( \hat{q} - \frac{\theta}{c_2} \) in each period.

With updating, the government has the ability to pick \( \tilde{q}_2 \) and \( \tilde{p}_2 \) after observing \( \theta \) and \( \eta \). In the case of prices, this is relatively straightforward because there is no behavioral link between periods and the government faces two distinct optimizations. In period one, \( \tilde{p}_1^U = c_1 \) leading to equilibrium quantities of \( \hat{q} - \frac{\theta}{c_2} \) as before, where superscript \( U \) indicates the government’s welfare maximizing solution to the updating policy problem. In period two, we have a first-order condition for the optimal updating policy \( \tilde{p}_2^U \) when all the uncertain outcomes are now known:

\[
\left( MB \left( h(\tilde{p}_2^U, \theta), \eta \right) - MC \left( h(\tilde{p}_2^U, \theta) \right) \right) h_p(\tilde{p}_2^U, \theta) = 0.
\]

Substituting and rearranging, we have

\[
c_2(\theta - \eta) + (b_2 + c_2)(\tilde{p}_2^U - c_1 - \theta) = 0,
\]

or

\[
\tilde{p}_2^U = c_1 + \frac{b_2\theta + c_2\eta}{b_2 + c_2}.
\]

This is the expression for the first-best, updated price in the second period. Note that in the second period, the regulator is able to achieve the first-best outcome as there is no longer an information asymmetry.

Optimal quantities are also straightforward. Let \( \tilde{q} = \tilde{q}_1 + \tilde{q}_2 \). From the previous discussion, we know that the firm chooses \( q_1 = q_2 = \tilde{q}/2 \). This is true regardless of how the government divides \( \tilde{q} \) between periods one and two. Therefore, the government’s problem amounts to figuring out the cumulative, two-period quantity policy \( \tilde{q}^U \) to maximize net benefits:

\[
\max_{\tilde{q}} \left[ 2B(\tilde{q}/2, \eta) - 2C(\tilde{q}/2, \theta) \right],
\]

68
where we let $\tilde{q}^U$ denote this optimized policy. Note that because the government can choose the second-period allocation after it observes the uncertain outcomes, it can determine the cumulative quantity target $\tilde{q}$ using the true values of $\theta$ and $\eta$, thereby circumventing the information asymmetry!

Working through the first-order condition

$$b_1 + \eta - b_2(\tilde{q}^U/2 - \hat{q}) - c_1 - \theta - c_2(\tilde{q}^U/2 - \hat{q}) = 0,$$

or, rearranging,

$$\tilde{q}^U/2 = \hat{q} + \frac{\eta - \theta}{b_2 + c_2}.$$ 

Letting $\tilde{q}^U_1$ be arbitrary, we have

$$\tilde{q}^U_2 = 2\hat{q} + 2 \frac{\eta - \theta}{b_2 + c_2} - \tilde{q}^U_1 = q^* + (q^* - \tilde{q}^U_1) \quad (3.8)$$

as the optimal, updated quantity policy.\textsuperscript{17} Intuitively, the second-period allocation is the first-best outcome for the second period $q^* = \hat{q} + \frac{\eta - \theta}{b_2 + c_2}$, plus a “top-up” $(q^* - \tilde{q}^U_1)$ to adjust the first-period allocation to the first-best level. This achieves the first-best outcome in both periods (in contrast to the price policy, which only achieves the first-best outcome in the second period). That is, knowing the firm will equalize marginal costs and divide the cumulative allocation across both periods, the regulator adjusts the quantity policy in the second period to accommodate the first-best outcome in both periods. This suggests that the comparative advantage for the updating policies will simply reflect the first-period loss under the price policy, which we now show.

\begin{footnotesize}
\textsuperscript{17} Because only the cumulative allocation matters, $\tilde{q}^U_1$ is not unique. There exists a continuum of optimal rational expectations equilibria, all of which have the property that $\tilde{q}^U_1 + \tilde{q}^U_2 = 2q^*$. This is because whatever the regulator chooses for $\tilde{q}^U_1$, setting $\tilde{q}^U_2 = 2q^* - \tilde{q}^U_1$ will achieve the cumulative allocation of $2q^*$.
\end{footnotesize}
3.3.3 Comparative Advantage

Without updating, the comparative advantage of prices versus quantities, $Δ^W$, is simply the original Weitzman result over two periods, or

$$Δ^W = 2Δ^O = \frac{σ^2}{c^2}(c_2 - b_2).$$

With updating, both policies achieve the first-best outcome in the second period. The comparative advantage stems from differences in the first-period outcome, namely the loss associated with the standard price instrument,

$$Δ^U = E[(B(h(\tilde{p}_1^U, \theta), \eta) - C(h(\tilde{p}_1^U, \theta), \theta)) - (B(\tilde{q}^U/2, \eta) - C(\tilde{q}^U/2, \theta))]$$

$$= -\left(σ^2_n + (b_2/c_2)^2σ^2_θ\right) \frac{2(b_2 + c_2)}{2(b_2 + c_2)}. \quad (3.9)$$

We provide an algebraic proof of (3.9) in the appendix, but a simple graphical proof is as follows. In the first period, the updating price policy is simply the standard price policy, $\tilde{p}_1^U = c_1$ and leads to the outcome at point $b$ in Figure 3.3. Meanwhile the updating quantity policy obtains the first best $q = \tilde{q}^U/2 = \hat{q} + \frac{\eta - \theta}{b_2 + c_2}$ at point $c$.

Whereas the quantity outcome is first best, we can see from the figure that the price policy leads to the shaded deadweight loss, $DWL$. This is a triangle where the “height” in the quantity dimension is $\frac{\theta}{c_2} + \frac{\eta - \theta}{b_2 + c_2} = \frac{\eta + (b_2/c_2)\theta}{b_2 + c_2}$. The “base” along the price dimension equals the height times $b_2 + c_2$. The area is then one-half base times height, or $\frac{(\eta + (b_2/c_2)\theta)^2}{2(b_2 + c_2)}$. Taking expectations of this area yields the expressed loss in (3.9).

---

18 The point where the $x$-axis intercepts the $y$-axis in Figure 3.3 should not be interpreted as the origin. It is zero for the $x$-axis but will be a negative number for the $y$-axis in the case of a bad. That is, this figure arises in the first quadrant for a good but the fourth quadrant for a bad.
3.3.4 Discussion of Key Assumptions

It is clear that the price policy is never preferred in this model. The quantity policy is first best in both periods while the price policy is generally only first best in the second period. In the special case where marginal benefits are flat ($b_2 = 0$) and there is no benefit uncertainty ($\sigma_\eta^2 = 0$), the relative advantage is zero. In such a case, prices achieve the first best. Thus, quantities with updating is weakly superior to prices. Quantities always achieves the first-best outcome and the welfare loss from price policies depends on both cost and benefit uncertainty.

This is quite different from the Weitzman result, without updating, where the preference between price or quantity controls depends on the difference between the marginal benefit and cost slopes. Moreover, both price and quantity controls are
second best in the Weitzman framework due to the information asymmetry faced by the government. Benefit uncertainty does not appear in the comparative advantage expression (3.5) because it equally affects the expected welfare loss from both prices and quantities.

The difference between the Weitzman result and our result is that there is an important interaction between policy updating and intertemporal trading. We have demonstrated that without updating and with intertemporal trading, the original Weitzman result remains in both periods. It is similarly easy to see that without intertemporal trading and with updating, the original Weitzman result remains in the first period (because the first period is just the Weitzman setup). But these two features together lead to a different result. In this setting, if firms have expectations about second-period policy updates before they make choices in the first period, the potential to trade across periods leads them to equalize current and expected future prices driven by the expected policy update. The result must be different. But must it achieve first-best?

Several additional implicit or explicit assumptions lead to the first-best outcome. First, we make a rational expectation assumption that firms are, in fact, expecting what the government ultimately does. This seems reasonable, as anything else should eventually lead to revised expectations. Second, we need to identify a quantity-policy updating rule for the government such that the expected next-period prices indeed achieve the first-best outcome. In the simple two-period model, we made that easy by assuming the periods were identical with a single common shock and no discounting. We will see in section 3.5 that more generally, with different shocks in different periods and more periods, there is still an updating rule that achieves the first-best in every period. Finally, we assume the government regularly updates policy based on societal net benefits and not some other objective. This is perhaps the most questionable assumption, to which we now turn.
3.4 Policy Updating with Noise

While improved cost-benefit information in future policy decisions can improve current-period regulatory outcomes under intertemporally tradable quantity controls, what if something other than improved information influences future policy? It is easy to imagine policy decisions being driven in part by special interests and more narrowly defined benefits rather than broad, societal costs and benefits. If we expect such changes to an intertemporally tradable quantity target in the future, we should be worried that these changes will be transmitted back to the present.

In particular, suppose that when the government updates its policy it acts as if marginal benefits in (3.4) are subject to an additive disturbance \( \epsilon \) reflecting uncertain special-interest political pressure rather than true information about societal benefits. One could also interpret this term as the Lagrangian multiplier on uncertain legal, political, or other constraints that bound the policy choice. Put simply, suppose the government now makes use of a “noisy” marginal benefit function given by

\[
MB_{\text{noisy}}(q, \eta, \epsilon) = MB(q, \eta) + \epsilon = b_1 + \eta - b_2(q - \bar{q}) + \epsilon
\]

that it wishes to equate with marginal costs in order to update its policy. The function \( MB \) without any subscript continues to refer to true, societal marginal benefits, while \( MB_{\text{noisy}} \) is the divergent, special-interest or constrained version of marginal benefits used to update policy.\(^{19}\) For simplicity, we assume there is no noise when policy is initially set for the first period, and \( \epsilon \) refers to the distortion introduced in period 1 along with the other true shocks to costs and benefits. In section 3.5 we extend this to an evolving noise process.

In adding this notion of political noise, we make the same information assumption about noise as we do about the true shocks to costs and benefits. Namely, the firm has

\(^{19}\) Note the regularity condition now becomes \( MB_{\text{noisy}}(0, \eta, \epsilon) > MC(0, \theta) \).
more up-to-date information about the evolution of political noise than is necessarily reflected in current policy. More specifically, the government sets policy for period 1, $\tilde{p}_1^N$ or $\tilde{q}_1^N$, before any political noise or shocks to costs and benefits occur. Here, superscript $N$ refers to the policy chosen in our setup with noise. By the time it acts in period 1, the firm knows the value of $\epsilon$ as well as $\theta$ and $\eta$ that will be used to update policies coming into effect in period 2. At the same time or soon after, the government sets $\tilde{p}_2^N$ or $\tilde{q}_2^N$ based on that same realization of $\theta$, $\eta$, and $\epsilon$.

Returning for a moment to our SO$_2$ trading program example, firms were well aware of evolving information, political forces, and legal opinions that were shaping the regulatory environment for 2010 and beyond. This awareness influenced pollution behavior and permit prices beginning as early as 2004. From 2004 to 2006, these changes were arguably driven by better information about welfare-improving, societal marginal benefits reflected in the initially proposed CAIR, which suggested such benefits were well in excess of $1600 per ton (Environmental Protection Agency 2005). During 2006-2009, the influences were arguably non-welfare-improving political forces and legal decisions, as the market price diverged from marginal benefits.

In our model terminology, consider the period 2004-2009 (and earlier years going back to 1995) as an elongated period 1. This policy was established under the 1990 Clean Air Act Amendments, which we would call period 0. During 2004-2009, firms expected the policy to be updated based on this period 1 information for 2010 and beyond, which we would call period 2.

Solving for the government’s optimal price policy is relatively straightforward because the choices each period are not linked by the firm. Choosing $\tilde{p}_1^N$ such that $E[MB_{\text{noisy}}(h(\tilde{p}_1^N, \theta), \eta, 0) - MC(h(\tilde{p}_1^N, \theta), \theta)] = 0$, we balance expected (noisy) marginal benefits and costs in period 1 as viewed in period 0. Note that the government is seeking to take expectations of $\theta$ and $\eta$ but not $\epsilon$. The value of $\epsilon$ — zero before the first period — enters directly, as the government’s current special interests
or current legal constraints are what matter for setting policy. This leads to

$$\tilde{p}_1^N = E[p^*]. \quad (3.11)$$

For period 2, uncertainty is revealed and noise is realized, so the government chooses $$\tilde{p}_2^N$$ such that $$MB_{\text{noisy}}(h(\tilde{p}_2^N, \theta), \eta, \epsilon) - MC(h(\tilde{p}_2^N, \theta), \theta) = 0$$, or

$$\tilde{p}_2^N = p^* + \frac{c_2\epsilon}{b_2 + c_2}. \quad (3.12)$$

The expression $$p^*$$ represents the first-best price outcome in terms of true aggregate net benefits,\(^\text{20}\) but the government’s objective now deviates from the goal of maximizing expected, aggregate net benefits. Using the firm’s reaction function (3.7), which is unchanged from our basic model, the equilibrium quantity outcomes are:

$$q_1 = h(\tilde{p}_1^N, \theta) = \hat{q} - \frac{\theta}{c_2}$$

$$= q^* - \frac{\eta + (b_2/c_2)\theta}{b_2 + c_2}$$

$$q_2 = h(\tilde{p}_2^N, \theta) = \hat{q} + \frac{\eta - \theta}{b_2 + c_2} + \frac{\epsilon}{b_2 + c_2}$$

$$= q^* + \frac{\epsilon}{b_2 + c_2}.$$

where $$q^*$$ is the first-best outcome in each period. These quantity outcomes are the same as those in our basic model plus the noise distortions in the second period.

Under a quantity policy with updating and intertemporal trading, we know the firm will simply set $$q_t = \frac{\tilde{q}_1^N + \tilde{q}_2^N}{2}$$ and what matters is this cumulative allocation as before. With the revealed values of $$\theta$$ and $$\eta$$ and the realized value of $$\epsilon$$, the government can choose $$\tilde{q}_2^N$$ based on an arbitrary choice of $$\tilde{q}_1^N$$ such that $$MB_{\text{noisy}} = MC$$ each

---

\(^{20}\) As noted in section 3.3.2, the first-best price is $$p^* = c_1 + \frac{b_2\theta + c_2\eta}{b_2 + c_2}.$$
period, or
\[
b_1 + \eta - b_2 \left( \frac{\tilde{q}_1^N + \tilde{q}_2^N}{2} - \tilde{q} \right) + \epsilon = c_1 + \theta + c_2 \left( \frac{\tilde{q}_1^N + \tilde{q}_2^N}{2} - \tilde{q} \right)
\]
\[
\tilde{q}_2^N = 2\tilde{q} + 2\frac{\eta - \theta}{b_2 + c_2} + 2\frac{\epsilon}{b_2 + c_2} + 2\frac{b_1 - c_1}{b_2 + c_2} - \tilde{q}_1^N
\]
\[
= q^* + \frac{\epsilon}{b_2 + c_2} + \left( q^* + \frac{\epsilon}{b_2 + c_2} - \tilde{q}_1^N \right),
\] (3.13)
which is again the same as the allocations in our basic model plus adjustments for the noise distortion. Given that firm behavior divides the cumulative allocation over both periods, we have
\[
q_1 = q_2 = \frac{\tilde{q}_1^N + \tilde{q}_2^N}{2} = q^* + \frac{\epsilon}{b_2 + c_2}
\] (3.14)
each period, which is the government’s optimal outcome given its own noisy view of marginal benefits, \( MB_{\text{noisy}}(q_t, \eta, \epsilon) \) including the \( \epsilon \) distortion. This outcome is clearly not first best from an aggregate welfare perspective.

What happens to our comparative advantage expression? In the second period, both policies produce the same outcome (namely \( q_2 = q^* + \frac{\epsilon}{b_2 + c_2} \)), so the comparative advantage derives solely from how the policies differ in the first period. In period 1, the noisy price policy obtains \( q_1 = q^* + \frac{\eta + (b_2/c_2)\theta}{b_2 + c_2} \) and the noisy quantity policy obtains \( q_1 = q^* + \frac{\epsilon}{b_2 + c_2} \). Both deviate from the first-best \( q^* \), and the comparative welfare advantage of a noisy price policy compared to a noisy quantity policy simply compares the relative losses arising from these two deviations,

\[
\Delta_N = \frac{\sigma_\epsilon^2 - (\sigma_\eta^2 + (b_2/c_2)^2\sigma_\theta^2)}{2(b_2 + c_2)},
\] (3.15)

Here, \( \Delta_N \) will be positive only if the variance from the added noise distortion, \( \epsilon \), under a quantity policy is larger than the variance from not getting the optimal price right...
\( p^* - E[p^*] \) under a price policy. If we think about the improved information about costs and benefits as a \textit{signal}, this amounts to asking whether the signal to noise ratio is greater than 1. Note that if \( b_2 = 0 \) and marginal benefits are flat, this simplifies to 

\[
\Delta^N = \frac{\sigma^2 - \sigma^2_\eta}{2(b_2 + c_2)} ,
\]

and the signal simply refers to improved information about benefits.

In our SO\(_2\) example, it would seem to be an open question whether the signal to noise ratio is \( \geq 1 \). The period from 2004 to 2006 was a large signal. The period from 2006 to 2009 was an equally large noise. Therefore, in this case the question of whether prices should be preferred to quantities remains open.

The important point is that when there is policy updating and intertemporal quantity trading, so long as firms have rational expectations, what matters for welfare is the signal to noise ratio associated with policy updates, not the relative slopes of marginal costs and benefits. This is what governs the relative advantage of price regulation compared to quantity regulation in this simple model. And that is because the key feature of quantity-based regulation in this context is not that it fixes the quantity, but that it allows expected future prices under the updated policy to govern the market price today.

### 3.5 Multiple Periods

We now extend our model and results to a multi-period setting. The purpose is to show that there are general price and quantity updating rules for the government that almost exactly replicate the previous two-period results comparing prices and quantities with noise (and without noise, if \( \epsilon = 0 \)). In expanding to multiple periods, we now assume cost and benefit uncertainty follow arbitrary AR(1) processes. That is, we keep the same expressions for marginal benefits and costs each period, (3.3)
and (3.4), but assume\textsuperscript{21}

\[ \eta_t = \rho \eta_{t-1} + \mu_t \]
\[ \theta_t = \rho \theta_{t-1} + \nu_t, \]

with \( E[\mu_t] = E[\nu_t] = 0 \). Similarly, we keep the same expression for noisy marginal benefits (3.10) but now assume the noise process follows a random walk,

\[ \epsilon_t = \epsilon_{t-1} + \omega_t, \]

with \( E[\omega_t] = 0 \). In contrast to the cost-benefit errors that may be mean reverting processes, we assume the government does not believe its current political position or constraints will revert back towards zero. Put another way, the current political situation is expected to persist.

We assume \( \eta_0 = \theta_0 = \epsilon_0 = 0 \) and \( \mu_t, \nu_t, \) and \( \omega_t \) are i.i.d. random errors for \( t > 0 \). We further assume a final period \( T \) after which \( \mu_t = \nu_t = \omega_t = 0 \) and \( q_t \) can be fixed forever at the desired rate by the government. This further assumption is not necessary; as discussed at the end of this section, we can obtain identical results (prior to period \( T \)) with an infinite horizon. However, the existence of a final period allows us to follow the intuition of the two-period model (and would be analogous to the idea that at some point the regulatory problem is fully solved).

The government’s policy choice for prices and quantities amounts to a dynamic programming problem each period, choosing today’s policy to maximize a given objective subject to future policies doing the same and firm profit-maximization. We assume that only information up to period \( t - 1 \) is used to set period \( t \) policy.

\textsuperscript{21} One may be concerned that the linear-quadratic approximation of marginal costs and marginal benefits may break down over long periods of time. It can be shown that our results all continue to hold if the \( b \) and \( c \) terms in the marginal cost benefit functions change arbitrarily over time. In particular, if those terms are additionally indexed by \( t \), the result continues to hold with appropriate adjustments to the allocation policy and trading ratios, but now the \( b_2 \) and \( c_2 \) values in the comparative advantage expression are similarly indexed by \( t \), and hence the comparative advantage varies over time.
Meanwhile, firms are able to use period $t$ information when they choose their (profit-maximizing) action in period $t$, including knowledge of how the government will behave in the future. This maintains the information structure assumed in the previous section as shown in Figure 3.4.

**Figure 3.4**: Timeline of events with updating in a $T$-period model

As before, firms facing quantity regulation with intertemporal trading have the flexibility to over- or under-comply each period. That is, the chosen quantity $q_t$ can be above or below the allocation $\tilde{q}_t$. With only two periods, we defined a bank that existed for a single period as the surplus or deficit given by $B = \tilde{q}_1 - q_1$, which was then applied to the second period allocation. With multiple periods, we might imagine a bank that accumulates based on $B_t = B_{t-1} + \tilde{q}_t - q_t$ until the final period $T$ when the bank in that period is applied to the final period allocation, similar to the two-period model. Here $B_t$ is the bank at the end of period $t$.

Instead, we introduce the one substantive difference between the multiple-period and two-period models: In the multiple-period model, we assume the existence of non-unitary trading ratios between periods under quantity regulation and assume the regulator is able to update those ratios alongside updates to the quantity allocation. Specifically, we modify the aforementioned bank accumulation rule in the following
way:

\[ B_t = \tilde{R}_t B_{t-1} + \tilde{q}_t - q_t, \]

The trading ratio is given by \( \tilde{R}_t \) and converts a one-unit deficit or surplus of period \( t - 1 \) quantities at the end of period \( t - 1 \) into those available at the beginning of period \( t \).

The idea of non-unitary trading ratios, which was previously proposed in Yates and Cronshaw (2001), is not as theoretical as it might seem. Such an approach has been proposed several times and used at least once. For example, the Clean Air Interstate Rule increased the stringency of the original Acid Rain Trading Program by effectively applying a trading ratio of 0.5 from 2009 to 2010 and 1/1.43 from 2014 to 2015.\(^{22}\) Similarly, the American Clean Energy and Security Act in the 111\(^{th} \) Congress (H.R. 2454, a.k.a. the “Waxman-Markey Bill”) proposed to exchange allowances in existing state programs for new federal allowances at a ratio that preserved the original value of the state allowances when they were first issued.\(^{23}\) Further, New Zealand’s emissions trading scheme has applied a “two-for-one” rule under which one permit allows for two tons of emissions; that rule is expected to change, effectively diminishing the future value of banked allowances. Indeed, New Zealand’s Climate Change Minister has said of the rule, “It was always a temporary measure...carbon must cost more than it does right now.”\(^{24}\) Finally, as China plans to transition from its regional carbon market pilots to a national-level program, it plans to allow permits from the pilots to be carried over to the national one at a discounted factor that depends on the degree of over-allocation in the pilots.\(^{25}\)

Having defined the multiple-period problem, we now need to find government

---

\(^{22}\) See 70 Fed. Reg. 91 (May 12, 2005), page 25291.

\(^{23}\) See Title VII, Subtitle B, Section 790.


policy rules that satisfy the first-order conditions for their objective function and mimic our simple two-period results. For a price policy, the dynamic programming problem simplifies to a series of one-period optimizations. We show in the appendix that setting
\[ \tilde{p}_t^M = E[p_t^* | t-1] + \frac{c_2 \epsilon_{t-1}}{b_2 + c_2} \]  
(3.16)
satisfies the government’s first-order condition,
\[ E[MB_{noisy}(h(\tilde{p}_t^M, \theta_t), \eta_{t-1}) | t-1] = E[MC(h(\tilde{p}_t^M, \theta_t), \theta_t) | t-1]. \]

The superscript $M$ indicates the multi-period result and $p_t^* = c_1 + \frac{b_2 \theta_t + c_2 \eta_t}{b_2 + c_2}$ is (as before) the first-best price outcome from a pure welfare perspective. That is, given the government’s position of setting policy one period ahead for period $t$, (3.16) represents its best effort to do so based on the available information $\eta_{t-1}$ and $\theta_{t-1}$ and subject to political noise $\epsilon_{t-1}$.

For a quantity policy, we also show in the appendix that if the government follows the allocation rule
\[ \tilde{q}_t^M = E[q_t^* | t-1] + \frac{\epsilon_{t-1}}{b_2 + c_2} + \tilde{R}_t \left( q_{t-1}^* + \frac{\epsilon_{t-1}}{b_2 + c_2} - \left( E[q_{t-1}^* | t-2] + \frac{\epsilon_{t-2}}{b_2 + c_2} \right) \right) \]  
(3.17)
and sets the trading ratio $\tilde{R}_t$ according to the rule
\[ \tilde{R}_t^M = \frac{p_t^* + (c_2/(b_2 + c_2)) \epsilon_{t-1}}{\beta (E[p_t^* | t-1] + (c_2/(b_2 + c_2)) \epsilon_{t-1})}, \]  
(3.18)
then the choice
\[ q_t = q_t^* + \frac{\epsilon_t}{b_2 + c_2} \]  
(3.19)
is a feasible choice for the firm each period. That is, it satisfies the terminal condition $B_T = 0$. In the above expressions, $\beta$ is the discount rate used by the firm and $q_t^* = \hat{q} + \frac{\eta_t - \theta_t}{b_2 + c_2}$ is (as before) the first-best quantity outcome. Note that both $\tilde{q}_t^M$ and
\( \tilde{R}_t^M \) only depend on \( t-1 \) information, and therefore meet the information structure that we assume.\(^{26}\)

After demonstrating (3.19) is feasible, we show this choice also satisfies the firm’s first-order condition for profit maximization, namely
\[
MC(q_t, \theta_t) = \beta \tilde{R}_{t+1} E[MC(q_{t+1}, \theta_{t+1})|t].
\]
Finally, we note that (3.19) is exactly the government’s choice to set \( MB_{\text{noisy}}(q_t, \eta_t, \epsilon_t) = MC(q_t, \theta_t) \) each period. That is, when acting with period \( t \) information (setting period \( t+1 \) policy), the government exactly achieves its period \( t \) optimum. Absent any noise (\( \epsilon_t = 0 \ \forall t \)), we have \( MB = MC \forall t \) and the government’s optimum exactly corresponds to the welfare-maximizing first-best outcome.

These multi-period results are logical extensions of our two-period results. The two-period price policy had different expressions for the first-period (3.11) and second-period (3.12) prices. However, these are just special cases of the finite-horizon, multi-period price outcome (3.16) corresponding to the first and last period. When \( t = 1 \), we have \( \epsilon_{t-1} = \epsilon_0 = 0 \) as we have assumed \( \epsilon_0 = 0 \). That is, \( \tilde{p}_1^M = E[p^*_1] \). This mimics the first-period price policy in the two-period model, which includes an expected first-best price term but no noise. And when \( t = T \), \( E[p^*_t|t-1] = p^*_T - (\mu_T + (b_2/c_2)\nu_T) = p^*_T \) as we have assumed \( \mu_t = \nu_t = 0 \) for \( t \geq T \). That is, \( \tilde{p}_T^M = p^*_T + (c_2\epsilon_T)/(c_2 + b_2) \). This mimics the second-period price policy in the two-period model, where there is a noise term but expectations are unnecessary because uncertainty is resolved. More generally, when \( T > 2 \) and \( 1 < t < T \), the price policy expression in (3.16) includes both the expected first-best price and the noise distortion.

Compared to the two-period quantity outcome (3.14), the multi-period quantity outcome (3.19) now has a time index on both the quantity optimum \( q^*_t \) and noise \( \epsilon_t \). It is otherwise identical. The government’s multi-period quantity rule (3.17) includes

\(^{26}\) The multi-period allocation rule has exactly the same form as the two-period rule in (3.13). That is, it includes the government’s best \( t-1 \) estimate of the desired allocation in period \( t \), namely \( E[q^*_t|t-1] + \epsilon_{t-1}/(b_2 + c_2) \). Plus, it includes a top-up for period \( t-1 \), based on the (now) correct desired allocation, \( q^*_t + \epsilon_{t-1}/(b_2 + c_2) \), minus the \( t-1 \) allocation previously provided based on period \( t-2 \) information, \( E[q^*_{t-1}|t-2] + \epsilon_{t-2}/(b_2 + c_2) \).
both a current-period allocation and a “top up” for the previous period, as does the two-period quantity rules (3.8) and (3.13).

The main difference is the introduction of the trading ratio $\tilde{R}_t^M$ defined in (3.18). This is necessary to allow the flexible control of prices over time, as one-for-one trading would necessarily imply expected prices rising at $1/\beta$ compared to current prices. However, trading ratios will be close to unity when the discount rate is small and when $\rho_\theta$ and $\rho_\eta$ are close to 1 (i.e., the cost and benefit processes are approximately random walks). In these cases, $\beta E[p_t^*|t-1] \approx p_{t-1}^*$, so $\tilde{R}_t^M \approx 1$ and non-unitary trading ratios may not be necessary to approximate the government’s optimal policy.

Turning to the comparative advantage of prices versus quantities in a multi-period setting, we have

$$\Delta^M = \frac{\sigma_\omega^2 - (\sigma_\mu^2 + (b_2/c_2)^2\sigma_\nu^2)}{2(b_2 + c_2)}.$$ (3.20)

for each period. Expression (3.20) closely resembles the two-period expression (3.15), with the variances of $\eta$, $\theta$, and $\epsilon$ replaced with those of the innovation terms $\mu_t$, $\nu_t$, and $\omega_t$. That is, in any period $t$ both price and quantity policies take into account all the information up through period $t-1$, $(\theta_{t-1}, \eta_{t-1}, \epsilon_{t-1})$. The difference is that, under the quantity policy, the government is able to encourage the firm to behave “optimally” in period $t$ based on period $t$ information and the promised update in period $t + 1$. This leads to a price policy loss determined by the cost/benefit innovation terms $\mu_t$ and $\nu_t$, but also a quantity policy loss determined by the added political noise $\omega_t$. This expression is derived in the appendix.

Note that while we assumed a terminal period $T$, we can easily extend this result to an infinite horizon by replacing with the terminal bank constraint with a government policy that places a finite limit on banking and borrowing. Intuitively, any price path that is a scalar multiple of the previous solution will continue to satisfy
the firm’s first-order condition. Previously, the terminal constraint on the bank narrowed this infinite set of choices to one: the government’s optimum. Without a terminal constraint on the bank, the firm would instead choose the least expensive path (e.g., $MC(q_t, \theta_t) = 0 \ \forall t$), with the bank accumulating any surplus or deficit. By establishing a finite limit on banking and borrowing, however, the government rules out any path except the one corresponding to their optimum. We discuss this further in the appendix.

3.6 Application to Climate Change Policy

Climate change policy has been lurking in the background as an important application of this paper and has motivated examples of stylized facts and assumptions. But what happens when we put quantitative estimates of climate change mitigation costs and benefits, and their associated uncertainties, into our expressions for comparative advantage?

Without political noise, our expression for $\Delta^M$ depends on the variance of cost and benefit shocks and the slopes of marginal costs and benefits. A standard approximation in the climate change literature is that marginal benefits are flat and $b_2 = 0$ (e.g., Newell and Pizer 2008). With that assumption, both $b_2$ and $\sigma^2\nu$ vanish from the expression and we only need estimates $\sigma^2\mu$ and $c_2$. Newell and Pizer (2003) provide an estimate of $c_2 = 1.6 \times 10^{-7}$$/ton^2$. We use the recent update to the U.S. Government’s Social Cost of Carbon (Interagency Working Group on the Social Cost of Carbon, United States Government 2013) to calibrate the error in marginal benefits. Namely, the central estimate changed from $24$/ton to $37$/ton, or by $13$/ton, in 3 years. Assuming $\eta$ follows a random walk, that gives 3-year variance of 169 ($$/ton)^2$
or \( \sigma_\mu^2 = 56 \text{ ($/ton)^2} \) for the 1-year variation. We then have

\[
\Delta M = \frac{\sigma_\omega^2 - (\sigma_\mu^2 + (b_2/c_2)^2\sigma_\nu^2)}{2(b_2 + c_2)}
\]

\[
= \frac{-\sigma_\mu^2}{2c_2} \quad \text{(ignoring noise and assuming } b_2 = 0) \]

\[
= \frac{-56 \text{ ($/ton)^2}}{2(1.6 \times 10^{-7}$$/\text{ton}^2)}
\]

\[
= -$175 \text{ million.}
\]

That is, if policy is set for each year \( t \) based on year \( t - 1 \) information (e.g., the policy is set one year in advance), the global loss from price policies is $175 million per year. Of course, policies are not set every year. The recent Paris agreement suggests policies (or at least nationally determined contributions) might be set every five years.\(^{27}\) In our model, the comparative advantage grows in magnitude with the time between policy updates, since the deadweight loss of a price instrument is repeated for every period in which prices are fixed at an inefficient level.\(^{28}\) With a one year interval between updates, the loss would be $175 million, but with a five year interval, the loss would be $2.4 billion in present value. Were policies updated every twenty years, which seems more in line with major U.S. policy adjustments (e.g., the 1990 Clean Air Amendments followed 20 years after the 1970 Clean Air Act; the 2010 adjustments to the SO\(_2\) regulations came 20 years after 1990 amendments), the loss would be $25.5 billion in present value for each 20 year period. Given a fifty year interval, the loss would be over $90 billion in present value.

\(^{27}\) See Article 4, paragraph 9.

\(^{28}\) Because we assume benefits follow a random walk, the variance in the second year will be twice the variance in the first year; the variance in the third year will be three times the variance in the first year (and so on). The sum of each year’s variance over \( t \) years is therefore equal to the \( t^{th} \) “triangular number”, \( t(t + 1)/2 \), times the variance in a single year. This sum must also be discounted; we use same the 3% discount rate used to generate the estimates for the social cost of carbon.
Of course, these estimates ignore the importance of political noise, $\sigma_\omega^2$, which favors price regulation. Unfortunately, such noise is inherently difficult to estimate. The 2015 Paris Agreement calls for limiting climate change warming to 2°C, or possibly 1.5°C. Meanwhile, the position of many climate skeptics would be that marginal benefits are close to zero (if not negative). While both of these positions might be viewed as the basis for a legitimate $MB$ estimate, they could also be viewed as a basis for political noise.

Rather than attempt to quantify $\sigma_\omega^2$ with such information, we make the observation that when marginal benefits are flat, the condition for price regulation to be preferred simplifies to $\sigma_\omega^2 > \sigma_\mu^2$. That is, prices are preferred if changes driven by political noise are larger than changes driven by true changes in estimated marginal benefits. Though a subjective judgment, it certainly seems plausible that this condition would be satisfied in practice.

### 3.7 Conclusion

After governments set their policies, new information often arises (or is revealed) about the benefits and costs of those policies. Much of the previous work comparing price and quantity regulation has focused on the importance of this information asymmetry in a one-period world or when policy remains fixed indefinitely. In this framework, Weitzman (1974) demonstrated that price-based regulation is preferred when marginal benefits are relatively flat compared to marginal costs and that the magnitude of the preference also depends on the variance of cost uncertainty.

We argue that governments eventually, if not regularly, update key regulatory policies. Moreover, quantity regulation is typically tradable over time. In this case, we show that what matters about quantity policies is not that they fix quantities but that they allow current prices to respond to expected future policy updates.
such policy updates seek to maximize welfare, we show that quantity regulation generally achieves the first-best welfare outcome and prices are never preferred. This highlights an important interaction between policy updates and intertemporal trading: Trading over time creates an arbitrage condition whereby expected future price changes will change prices today. This can be used by the regulator to overcome the information asymmetry. Such near-term anticipatory influences cannot arise under price-based regulation, which will only change prices at the moment the future policy update comes into force.

It is widely recognized that expectations of all types of future policy can have immediate effects on investment behavior through future expected prices. This result highlights that, under intertemporally tradable quantity regulation, such expectations can also affect contemporaneous prices through a no arbitrage condition, with immediate impacts on production and consumption behavior.

When we consider the possibility of noisy policymaking, where something other than welfare maximization drives the policy updates, a trade-off between price and quantity regulation re-emerges. Quantity controls with intertemporal trading and policy updates allow both new information about true costs and benefits and new noise to drive current prices. The relative advantage expression then depends on the difference between these variances. When marginal benefits are flat (as in the case of carbon dioxide emissions and global climate change), the sign of the expression simplifies to depend only on the difference between the added noise variance and the added marginal benefit variance. This is quite different from the original Weitzman result and most of the existing literature.

These results provide an important complement to that existing literature. Namely, the traditional perspective makes sense when policy updates are sufficiently far off, unspecific, or unforeseen so as to make them irrelevant to market participants. However, as the possibility of an update becomes more proximate and specific, these
newer results should be considered.

There are other caveats to our results. The specific expressions for comparative advantage assume the use of possibly non-unitary trading ratios between periods that depend on the nature of cost and benefit shocks and discounting. There is only limited experience with such ratios in practice. Price and quantity policies are also rarely enacted in isolation; most exist alongside myriad other regulations. California, for example, has at least four major regulations operating in tandem with its emission trading program (Borenstein et al. 2015). We have not considered how such policies affect our results.

Further, there are other types of policy imperfections aside from the political noise as we have modeled it. For example, the change in allowance prices caused by changing expectations redistributes value. Savvy, politically active firms may intentionally pressure regulators to announce policy changes in attempts to manipulate the value of their allowance holdings, making the noise itself endogenous. We leave this as a topic for future research.

Finally, we assume there is a significant lag in policy updates between when the change is well-established and when it occurs, and that lag is the same for both price and quantity policies. This is relatively standard in regulatory policy, including quantity-based policies, where regulated firms are given considerable lead time and the process itself involves many steps. Tax policy changes, however, typically happen quickly to prevent avoidance behavior. It is unclear whether updates to price-based regulation would follow more of a tax policy precedent, or a regulatory policy precedent. A difference in the updating process and timing between price and quantity regulation would require further analysis.

Regardless of these caveats, important intuition remains. The traditional (economic) debate over price versus quantity regulation has emphasized relative slopes. When we consider the reality of policy updates over time and intertemporally trad-
able quantity regulation — indeed, when we look at the historic behavior under existing programs — it is clear that changing price expectations often matter. That may or may not favor prices versus quantities, but it does require additional consideration.
Chapter 4

Trophy Hunting vs. Manufacturing Energy: The Price-Responsiveness of Shale Gas (Co-authors: Richard G. Newell and Ashley B. Vissing)

Note: this work was co-authored with Richard Newell and Ashley Vissing. Prest’s contribution involved collecting and cleaning the data, developing and estimating econometric models, devising and computing the simulation models, and writing the initial draft. Developing the hazard models and editing the drafts were performed collaboratively.

4.1 Introduction

Over the past decade, technological developments have driven increased natural gas and oil extraction by opening access to resources stored in shale and other “tight” formations. These unconventional technologies and resources have allowed drillers to extract from significantly larger subsurface acreage using fewer wellbores and with much higher production per well. The combination of hydraulic fracturing and
horizontal drilling techniques has underpinned this unconventional supply and the resulting shale gas boom.\(^1\) Supported initially by high natural gas prices during much of the 2000s, the United States has experienced significant increases in natural gas and oil production (see, e.g., Hausman and Kellogg 2015; Kilian 2016 and Kilian 2017a).

The shale revolution has fundamentally changed how gas and oil are produced in the United States. In the words of one industry expert,\(^2\) conventional oil and gas investments resemble high-risk/high-reward, “big game trophy hunting,” which involves drilling many dry holes in search of a few highly productive ones. This stands in stark contrast to modern unconventional extraction from shale, which is commonly said to resemble a “manufacturing process” in that operators have much more flexible and certain control over their production levels.

Multiple features of unconventional gas lead to this flexibility. First, industry operators have long known about the location and scale of shale resources, but the technology did not exist to cost-effectively extract those resources in significant amounts until the past decade. As one industry analyst stated, “we knew the shale formations were there but we didn’t have the technology to extract from them.”\(^3\) In particular, shale resources tend to be more uniformly distributed geographically, making shale gas deposits easier to identify, reducing uncertainty about well productivity and leading to fewer dry holes. Second, shale resources produce a much larger amount of resources quickly, suggesting a tighter relationship between drilling effort and realized production. Altogether, experts have suggested that these factors make

---

1 Conventional wells tap porous and permeable resources that flow to the wellhead naturally once the well is drilled. Unconventional wells require hydraulic fracturing, where tight shale resources are artificially stimulated with high-pressure water causing fissures that are propped open to allow the gas and oil to flow to the wellhead. The process requires more time and at a higher cost to complete the well.

2 Rob Jacobs, personal communication.

unconventional gas and oil more responsive to prices.⁴

To the extent that unconventional gas is more price responsive, the shale boom has not only shifted the supply curve but “flattened” it—that is, a larger increment of gas is expected to be produced in response to an exogenous increase in gas prices. A flatter supply curve in turn should reduce gas price volatility, ceteris paribus, as illustrated in Figure 4.1. That figure shows the effects of an approximately 3-fold horizontal scaling of the supply curve, thereby reducing price volatility due to demand shocks.⁵,⁶

In practice, natural gas prices have become significantly less volatile following the boom in shale gas, compared to the early 2000s. It is possible that the increased price responsiveness of U.S. natural gas supply could explain this decline in volatility as depicted in Figure 4.1. To the extent that unconventional gas is responsible for this diminished volatility, continuation would help reduce uncertainty for policymakers and businesses considering investments that are highly sensitive to gas prices. For example, increasing reliance on natural gas-fired generation increases the exposure of utilities to natural gas prices, which have historically been more volatile than coal prices. The economic benefits of investments in LNG export infrastructure also

---


⁵ This conclusion rests on several assumptions, each of which is plausible. First, it assumes that the gas price volatility is primarily due to demand shocks (e.g., cyclical fluctuations in temperature, affecting heating and cooling demand), rather than supply shocks (e.g., offshore well shut-ins due to extreme hurricanes like Katrina and Rita). In practice, gas price volatility tends to track temperature variation, supporting this assumption. Davis and Kilian (2011a) also shows the relative importance of demand as a key driver of natural gas prices. Second, it assumes that the variances of the shocks to the supply and demand schedules have not also risen. If anything, supply shocks are likely smaller due to the shale revolution for two reasons: the lesser uncertainty about shale wells’ productivity and the lower vulnerability to hurricanes, which primarily affect offshore gas production.

⁶ This figure illustrates a distinction between price responsiveness and elasticity. The illustrative unconventional supply curve has everywhere a larger price response than the conventional (specifically, it has one third the slope). But it is not everywhere more elastic. Nevertheless, it still leads to lower price volatility. Hence a change in price response through an increase in productivity could lead to lower price volatility, even with an unchanged elasticity.
depend on stable natural gas prices, as do the benefits of domestic investments in energy-intensive manufacturing and chemical production.

Given the differences in geology and changes to drilling technology, this research seeks to disentangle the differences in operators’ price-responsive behavior between conventional and unconventional gas. We consider different aspects of the production process to assess the differences between unconventional and conventional wells at each stage of the supply function, beginning with the decision to drill the well, complete the well, and produce gas over time. We estimate these relationships using econometric models that appropriately describe each stage of well development and execute the analysis using detailed data from approximately 62,000 gas wells drilled in Texas between January 2000 and September 2015.

We find that the important margin for production responses to price changes is drilling investment, whereas neither production from existing wells nor the time from drilling to first production respond strongly to price changes. We estimate a cumulative gas drilling response to prices of approximately 0.9 and find no clear

Figure 4.1: Illustration of the Effect of Flatter Supply on Price Volatility
evidence that the elasticity is different for unconventional compared to conventional
gas drilling. While unconventional wells take somewhat longer to reach production,
they produce much more gas per well—about 3 times as much—compared to conven-
tional wells and have less risk associated with variation in well productivity. This
faster flow rate per well turns out to be the primary margin by which aggregate sup-
ply from unconventional gas production is more price responsive than conventional
production. While this study focuses on natural gas drilling in Texas, our analysis
of gas wells nationwide finds that the productivity of Texas wells is approximately
representative on average of gas wells nationally.7

We use these econometric results to conduct simulations of the effect of exoge-
 nous shocks to gas prices on wells drilled and gas supplied over time, providing a
time-varying price response. We find that on a per-well basis, unconventional gas
production responds more strongly to price changes (about 3-fold in the long-run)
than conventional production does. Accounting for changes to the level and compos-
sition of drilling activity, the responsiveness of natural gas supply is about 3 times
as price responsive during the “shale era” of 2010-2015 compared to the “pre-shale
era” of 2000-2005. These are the first econometric estimates to isolate the differences
in supply response between shale gas and conventional gas.

Our results also highlight the inertia in oil and gas supply dynamics owing to
the fact that wells produce for many years after being drilled, meaning that gas
production today depends in part on price changes many years ago. This implies
that simply regressing gas production on current and/or a few lags of gas prices is
likely to miss these short- relative to long-run dynamics.

7 In particular, the average productivity of Texas gas wells and gas wells nationally differs by less
than 10 percent, for both unconventional and conventional wells. In Texas, we find unconventional
gas wells to be about 2.7 times as productive as conventional wells on average, compared to a ratio
of 3.2 nationwide.
4.2 Literature

The paper spans several literatures related to the economics of the oil and natural gas industry and more generally to non-renewable resource extraction. We contribute to a nascent but growing area of research on the effects of greater accessibility to shale gas (Joskow 2013). The study also contributes to understanding of price formation in fossil fuel markets (e.g., Hamilton 2009; Kilian 2009; Baumeister and Kilian 2016a). Specifically, we analyze the separate phases of energy development and isolate the industry decisions that are responsive to changes in natural gas and oil prices, effectively estimating a supply response in a much more disaggregated manner than is typical in past work.

Our paper empirically tests the degree to which natural gas production decisions are sensitive to shocks in prices and whether there are differential effects for conventional versus unconventional resources and technologies. Using more disaggregated, well-level data applied to gas (rather than oil), we find evidence that the quantity produced from already-producing wells is not price responsive, consistent with Anderson et al. (2014).

The demand and supply elasticity literature often compares short and long-run results, typically finding that gas and oil supply elasticities are inconsequential once wells have been drilled and that supply is less responsive in the short run than in the long run. Indeed, it is common in the empirical literature using structural vector autoregressions to assume a short-run price elasticity of zero (see Kilian 2009; Kilian and Murphy 2012 and Kilian and Murphy 2014, although Baumeister and Hamilton (2015) find evidence for a small, positive short-run supply elasticity). Papers analyzing oil extraction elasticities include Griffin (1985); Hogan (1989); Jones (1990); Dahl and Yücel (1991); Ramcharran (2002); and Güntner (2014). There is very little published econometric evidence on natural gas supply elasticities (Erickson and
Spann 1971; Dahl 1992; Krichene 2002), much of which dates to before the recent shale gas revolution.

Our paper is specifically focused on estimating the heterogeneous responses for unconventional compared to conventional drilling decisions, centered on the recent shale gas boom occurring in states such as Texas, Pennsylvania, Louisiana, and Arkansas. Hausman and Kellogg (2015) estimates aggregated supply and demand elasticities to capture welfare impacts of the shale boom across industrial sectors and producers, and heterogeneous effects across space. They estimate a drilling elasticity of 0.81 using aggregate data. Mason and Roberts (2016) also estimates drilling and intra-well gas supply elasticities using data on wells in Wyoming, find drilling elasticities of 0.61 to 0.82. This study estimates a drilling response of 0.89, similar to those two studies despite using different methods and datasets.

This study is differentiated from those previous studies as follows. First, we separately estimate price responsiveness at multiple stages of production, including the “time to completion” margin, whereas past studies have not. Second, we separately estimate supply elasticities for conventional versus unconventional sources, which past studies have not. Third, we provide an integrated model of the multiple stages of production to simulate the effect of a price shock on drilling, completion, and production over time.

Other relevant studies include Rao (2013), Metcalf (2017), and Smith and Lee (2017). Rao (2013) uses well-level microdata to estimate production responses to windfall profits taxes imposed in California in the 1980s. Metcalf (2017) estimates the impacts of removing certain tax preferences for U.S. oil and gas drilling, which involves modeling oil and gas supply responses. Smith and Lee (2017) estimates the price responsiveness of shale oil reserves (as opposed to production). This study complements and contributes to this line of research by estimating supply elasticities at multiple stages of the production process.
Our study also contributes to understanding decline curves for unconventional wells, which has important implications for the inertia and cyclicality inherent in gas and oil markets. Decline curves from conventional wells are commonly modeled using the “Arps equation,” which nests exponential, hyperbolic, and harmonic decline curves. By contrast, Patzek et al. (2013) argue that unconventional gas wells in the Barnett shale formation follow a fundamentally different functional form: proportional to 1 divided by the square root of time for the first few years, and then exponential after that. This same functional form is also used in Browning et al. (2014), among other studies. While we explore this functional form in this paper, we ultimately focus on a non-parametric approach to decline paths.

4.3 Industry Background and Data

4.3.1 Industry Structure

Unconventional drilling describes the technological combination of more advanced hydraulic fracturing and horizontal drilling techniques that are used to extract natural gas and oil from tight-shale formations. Unconventional gas and oil reserves are stored in these tight-shale formations where the resources are difficult to extract unless the shale is artificially stimulated using a technique like hydraulic fracturing.

Further, large quantities of shale resources are located beneath more densely populated regions of the country, and these resources can now be accessed through horizontal drilling techniques, with laterals extending from the vertical well for thousands of feet in any horizontal direction. Combined with advancements in seismic and surveying technologies, these extraction methods increase the accessibility to otherwise inaccessible natural gas and oil.

As a result, unconventional gas and oil production differs from conventional methods in several ways across the stages of development, which are described in the
following subsections. In particular, the process can be broken down into the following stages: leasing and permitting; spudding a well (i.e., commencing drilling); well completion and stimulation; and production over time.

**Leasing and Permitting**

Before obtaining a permit to drill a well, a firm signs leases with the mineral rights holders of acreage from which the firm wants to extract natural gas and/or oil, and these mineral rights holders may be private landowners or government entities. During the term of the lease, the lessor (landowner) retains the ownership and use of the surface estate while the lessee (firm) has the right to extract and sell the resources stored in the sub-surface mineral estate. In return for the right to extract, the lessor is paid a lump-sum bonus, land rental payment, and royalty on the value of the extracted resources.

Once the mineral rights are leased, firms apply for a permit to drill a well from the relevant state regulatory agency. While we investigated estimating the permitting stage explicitly, this stage did not add significantly to the analysis so we did not include it in our model. Permitting tends to be quick and low-cost, in contrast to drilling and completion which require significantly larger investments. For further published research on the state and federal land leasing decisions see Porter (1995); Hendricks et al. (2003), and Fitzgerald (2010), and for privately owned mineral leasing see Vissing (2015).

**Drilling**

Once mineral acreage is leased and a well is permitted, drilling can commence. While drilling unconventional wells results in greater oil and natural gas production per well, it is also more technically challenging and expensive than conventional onshore extraction. An unconventional well is drilled both vertically and then horizontally
as compared to a conventional well that is drilled in only vertically. The vertical segments typically reach 4,000 to 13,000 feet in depth and the horizontal segments typically extend 2,000 to 7,000 feet. The horizontal segment allows firms to extract more natural gas from a larger acreage using a single wellbore. The date at which drilling begins is called the “spud date.” The firm that owns the well typically contracts the drilling task out to specialized oilfield service companies (e.g., Schlumberger and Halliburton) that charge daily rates to rent their drilling rigs and services. These rig rates are on the order of $15,000 per day.\(^8\)

**Well Completion and Stimulation**

Following the drilling phase, a well must be completed. Completion primarily involves casing, perforating, and possibly stimulating the well. A casing is a large diameter pipe cemented into the wellbore, and it prevents the well from caving in and protects groundwater from contamination. The horizontal part of the casing in the shale formation is perforated using perforating guns containing explosive charges so that gas and/or oil can flow into the casing and up to the surface.

Permeable conventional reservoirs have natural pressure that causes the resource to flow easily to the surface without any additional work. However, the impermeability of tight-shale formations requires artificially stimulating the well through large-scale hydraulic fracturing techniques. This process injects millions of gallons of water mixed with sand and chemicals into the well at a high pressure. The pressure causes the rock to fracture, generating fissures in the rock that are propped open by the sand in the fracturing fluid. Once the fracturing fluid returns to the surface, the newly propped fissures allow the gas and oil to flow more easily up to the wellhead.

\(^8\) However, this is a small piece of the total cost of developing a well. For example, according to one industry expert, the overall cost of developing a well has been on the order of $4-12 million, whereas payments to drillers may make up only $0.5 million of this (30 days at $15,000 per day).
Production over Time

Once a well is completed, it often produces oil and natural gas for many years or even several decades. The largest flow of production is realized in early periods because of high reservoir pressure. Production from existing wells involves little marginal cost, so firms are incentivized to produce quickly to recover drilling and completion costs. The flow rate of oil and natural gas is therefore a direct function of the quantity of resources remaining in the ground. As more hydrocarbons are extracted, there is less natural pressure pushing what remains to the surface, reducing the flow rate.

As a result, the flow rate of a typical well decreases quickly. In our data, gas production typically decreases more than 60 percent from its peak after only one year. In addition, in the first year of production, conventional and unconventional gas wells realize 35 percent and 34 percent, respectively, of their total gas production. In the first five years, they realize 78 percent and 77 percent of total production, respectively. Further, because the marginal cost of production from an existing well is fairly low, many wells continue to produce even after their production capacity has fallen to near zero. Firms also have alternative extraction methods to increase the flow rate in later periods, including reservoir stimulation methods and pumps (for oil only). However, our data does not directly indicate whether firms have employed these secondary extraction techniques to increase flow rates. Another factor affecting a well’s productivity is well design, such as the length of laterals, fracturing inputs (for shale), perforation, completion design, and so on.

4.3.2 Data Sources

We use well-level data aggregated by Drillinginfo, a company that provides information services on upstream oil and natural gas activity. We focus on wells drilled in
the state of Texas because of the superior quality of the data from that state. The data for other regions tend to be inferior because certain variables are unavailable, defined differently, or reported less frequently. For example, in New Mexico, drilling direction is not reliably tracked; in Pennsylvania, production is reported at the annual level, compared to monthly in Texas.

9 The data for other regions tend to be inferior because certain variables are unavailable, defined differently, or reported less frequently. For example, in New Mexico, drilling direction is not reliably tracked; in Pennsylvania, production is reported at the annual level, compared to monthly in Texas.

10 http://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_FGW_mmcf_a.htm.

11 One reader noted that due to a one-year period during which firms need not report data to the Texas Railroad Commission, the data available as of June 2016 may not be 100 percent complete for wells drilled after June 2015. In all of our analysis, we only use wells drilled through September 2015. All of our results are robust to restricting the sample to only wells drilled through June 2015.
We focus on natural gas wells, ignoring oil wells. We also compute the length of horizontal well “laterals” using the geodesic distance between the well’s surface hole and bottom hole.

We categorize a well as unconventional if it is drilled in a low-permeability or shale reservoir using horizontal or directional drilling techniques. Reservoirs were classified using a dataset provided by the Energy Information Administration (EIA) that labels field and reservoir names as either “conventional,” “low permeability,” or “shale.” Correspondingly, a conventional well is defined as a well that is drilled vertically into a conventional reservoir.

Unless otherwise noted, we use the simple average of the next 12 months of futures

12 These occasional inconsistencies are primarily due to the way Drillinginfo updates data on re-entered wells and likely represent data errors.

13 The definition of a “gas” well versus an “oil” well is determined by engineering-based regulatory standards for well spacing. In our data, observable well characteristics are consistent with the reported “oil” or “gas” label. For example, on average gas wells obtain 94 percent of their cumulative energy production (in MMBTU) from gas, compared to only 23 percent for oil wells. Further, 72 percent of gas wells obtain 95 percent or more of their output from gas, compared to only 0.5 percent of oil wells. Finally, very few unconventional oil wells are reported before 2010, whereas there is a strong and steady upward trend in unconventional gas wells during 2005-2010 (see Figure 4.4).

14 Oil wells also often co-produce natural gas, as well, accounting for approximately 20 percent of total gas production (both nationwide, and in Texas specifically). We focus on the 80 percent of gas production that comes from gas wells. Other work has suggested that the drilling of oil wells does not respond to natural gas prices, but is instead driven by oil prices. Therefore, the major price response to gas prices comes from gas wells, which is our focus. This highlights the importance of distinguishing between types of wells for analyzing gas supply.

15 Approximately half of the wells matched perfectly to EIA’s dataset. The remaining wells were matched using the closest reservoir name by Levenshtein distance, the number of character changes required to obtain a perfect match. 78 percent of the wells match perfectly after changing no more than 1 character; 93 percent match perfectly after changing no more than 4 characters. Many such close matches are either misspellings, abbreviations, or spelling variants (e.g., “Barnett shell,” or “Eagle Ford” versus “Eagleford”) or alternatively include extra information (e.g., “Helms Barnett shale”). Inspection of the results suggests the match performs well. For example, very few wells classified as unconventional appear before 2000.

16 “Mixed” wells (i.e., a vertical well in a shale reservoir, or a horizontal/directional well in a conventional reservoir) are dropped from the analysis, although there are relatively few of these. Of the wells designated as unconventional, 78 percent are in shale reservoirs, 15 percent are in horizontally-drilled low-permeability reservoirs, and the remaining 7 percent are in directionally-drilled low-permeability reservoirs.
prices for natural gas price (Henry Hub) and oil (WTI). Each price is the average of daily prices collected from Bloomberg and adjusted to 2014 dollars using the CPI All Urban Consumer (All Items) index.\textsuperscript{17}

\textbf{4.3.3 Data Description}

The cleaned dataset includes approximately 62,000 gas wells drilled between 2000 and 2015. For unconventional wells, we only include wells drilled in 2005 or later, as this is the period when the shale gas revolution began in earnest. Figure 4.2 shows a map illustrating the location of the wells in our data along with depictions of selected shale plays. Table 4.1 reports the summary statistics for conventional and unconventional gas wells,\textsuperscript{18} along with summary statistics for our price data.

Unconventional wells are much more productive than their conventional counterparts. On average, an unconventional gas well in our data produced nearly 70,000 thousand cubic feet (mcf) of natural gas in its first full month,\textsuperscript{19} compared to approximately 30,000 mcf from a conventional gas well, meaning on average over this sample period (2000-2015 for conventional and 2005-2015 for unconventional), unconventional wells are 2.3 times as productive.\textsuperscript{20} However, this average masks trends in productivity; per-well productivity has been rising substantially for unconventional wells and falling slightly for conventional wells.\textsuperscript{21} Over the 2010-2014 period, the average initial production for unconventional wells was 80,000 mcf per month (not shown in table), 2.7 times as productive as the average conventional well (30,000 mcf per month).

\textsuperscript{17} In all cases, we use the realized values of the CPI index.

\textsuperscript{18} Decline rates represent decline from peak production, which is not necessarily the first full month. Means, medians, and standard deviations are taken over relevant subsets of the data. For example, wells that produced no oil are excluded from the calculation regarding oil decline rates.

\textsuperscript{19} Initial production is measured as production during its first full month of production, meaning the second calendar month during which production is reported. It is standard to focus on the second month because a well is typically only producing for a fraction of its first calendar month.

\textsuperscript{20} The t-statistic for this difference is -70, so $p < 0.001$.

\textsuperscript{21} See appendix for graphs of average well productivities over time.
According to industry sources, the causes of this rising productivity are primarily technological, including longer and better targeted laterals and better completion design.

Figure 4.3 shows the distribution of initial production on a log scale for both mcf over 2000-2015, as shown in Table 4.1).\textsuperscript{22} According to industry sources,\textsuperscript{23} the causes of this rising productivity are primarily technological, including longer and better targeted laterals and better completion design.

\textsuperscript{22} The t-statistic for this difference is -74, so $p < 0.001$. In this calculation, we focus on the more recent 2010-2014 period for unconventional wells to account for the upward trend in their productivity (while excluding 2015 to exclude the impacts of the sharp decline in oil prices that year). We use the longer sample of 2000-2015 for conventional wells because relatively few conventional wells were drilled after 2010, and focusing on that time period could raise selection concerns due to the crowding out from unconventional drilling. Nonetheless, the relative productivities are not strongly affected by this choice of time period for conventional wells. Focusing only on wells (of both types) from 2010-2015, the unconventional productivity is still approximately 3 times as large as conventional productivity.

\textsuperscript{23} Rob Jacobs, personal communication
Table 4.1: Summary Statistics

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Conventional</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Std. Dev.</td>
<td>Mean</td>
<td>Median</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Well Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Gas Production (first full month, mcf)</td>
<td>30,261</td>
<td>14,214</td>
<td>55,055</td>
<td>68,244</td>
<td>49,918</td>
<td>68,183</td>
</tr>
<tr>
<td>First 12 Months’ Total Gas Production (mcf)</td>
<td>217,060</td>
<td>106,619</td>
<td>397,392</td>
<td>493,399</td>
<td>371,878</td>
<td>463,673</td>
</tr>
<tr>
<td>Gas 3-Month Decline Rate (%)</td>
<td>47.9</td>
<td>46.8</td>
<td>23.2</td>
<td>43.4</td>
<td>40.6</td>
<td>20.5</td>
</tr>
<tr>
<td>Gas 12-Month Decline Rate (%)</td>
<td>70.8</td>
<td>72.9</td>
<td>19.7</td>
<td>68.3</td>
<td>68.1</td>
<td>16.6</td>
</tr>
<tr>
<td>Gas 24-Month Decline Rate (%)</td>
<td>80.5</td>
<td>83.0</td>
<td>16.1</td>
<td>78.8</td>
<td>79.4</td>
<td>13.6</td>
</tr>
<tr>
<td>Initial Oil Production (first full month, barrels)</td>
<td>344</td>
<td>0</td>
<td>1,347</td>
<td>2,245</td>
<td>0</td>
<td>5,050</td>
</tr>
<tr>
<td>First 12 Months’ Total Oil Production (barrels)</td>
<td>2,311</td>
<td>150</td>
<td>8,770</td>
<td>14,503</td>
<td>253</td>
<td>32,140</td>
</tr>
<tr>
<td>Oil 3-Month Decline Rate (%)</td>
<td>70.3</td>
<td>72.7</td>
<td>26.6</td>
<td>64.6</td>
<td>63.7</td>
<td>26.8</td>
</tr>
<tr>
<td>Oil 12-Month Decline Rate (%)</td>
<td>86.0</td>
<td>92.4</td>
<td>17.7</td>
<td>84.3</td>
<td>87.4</td>
<td>16.0</td>
</tr>
<tr>
<td>Oil 24-Month Decline Rate (%)</td>
<td>91.5</td>
<td>97.9</td>
<td>13.1</td>
<td>91.1</td>
<td>94.3</td>
<td>11.2</td>
</tr>
<tr>
<td>Horizontal Well Length (ft)</td>
<td>8,963</td>
<td>9,250</td>
<td>3,287</td>
<td>12,414</td>
<td>11,780</td>
<td>3,066</td>
</tr>
<tr>
<td>Total Vertical Depth (ft)</td>
<td>2.53</td>
<td>2.00</td>
<td>3.08</td>
<td>4.73</td>
<td>4.00</td>
<td>3.64</td>
</tr>
<tr>
<td>Months Between Spud Date and First Production</td>
<td>36,093</td>
<td>26,017</td>
<td>1,836</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Wells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price Data (Monthly, 2000-2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henry Hub Natural Gas Price, Prompt Month ($/MMBtu)</td>
<td>$5.95</td>
<td>$5.02</td>
<td>$2.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henry Hub Natural Gas Price, 12-Month Future ($/MMBtu)</td>
<td>$6.31</td>
<td>$5.63</td>
<td>$2.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTI Oil Price – Prompt Month Future ($/barrel)</td>
<td>$70.90</td>
<td>$73.12</td>
<td>$27.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTI Oil Price – 12-Month Future ($/barrel)</td>
<td>$71.15</td>
<td>$76.13</td>
<td>$27.72</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Figure 4.3: Distributions of Initial Gas Production (IP) by Well Type, Log Scale

Sources: Authors’ calculations based on data from Drillinginfo and EIA.
conventional wells compared to conventional ones. Figure 4.3 also illustrates that unconventional wells are less likely to be “dry”, as exhibited by the relatively smaller mass in the lowest productivity bin for unconventional wells. These differences are likely due to the geological features of shale resources discussed above.

On average, we do not observe a substantial difference in decline rates between unconventional and conventional gas wells—somewhat contrary to conventional wisdom. Decline rates measure the rate of change in production levels for a given well beginning with the peak period, typically but not always the first full calendar month of production. Unconventional gas wells extract more natural gas in earlier periods because they are much larger on average. Unconventional gas wells also have more consistent decline rates than conventional wells, as demonstrated by the lower standard deviations for the decline rates for unconventional wells in Table 4.1.

According to industry participants, unconventional wells are more expensive to drill and complete than conventional wells, so the gain in physical well productivity also comes with higher costs. The primary costs associated with developing an oil or gas well are associated with renting rigs during the drilling and completion stages. These expenses depend on the amount of time the rigs are working on site, which depends in turn on the depth of the well and the length of its laterals, also summarized in Table 4.1. Average lateral lengths have been rising at a remarkably linear pace over that period (not shown), from about 2,000 feet in 2005 to about 6,000 feet in 2015, with an average of about 4,000 feet over the full sample period. Further, in section 4.4.3 we consider the length of time between spudding and first production, which is greater for unconventional wells due to the extra steps required to hydraulically fracture the well: unconventional gas wells begin producing within

---

24 One could further explore such variation by distinguishing between infill wells, which involve developing established resources, versus wildcat wells, which involve exploring for new deposits. A reader suggested that such wells may have different characteristics and supply responses. Exploring those is beyond the scope of this paper. In effect, our analysis averages over these different types of wells to estimate an overall supply response, which is the primary goal of this study.
4.7 months, on average, compared to 2.5 months for conventional gas wells.

Figure 4.4 describes the number of new spuds each quarter of our data from 2000 to 2015 by well type (conventional versus unconventional gas wells), and the spud counts are plotted along with natural gas and oil prices. The figure reveals that the wells we have classified as unconventional are indeed a recent phenomenon, supporting our classification method. Following the massive increase in unconventional drilling and collapse in gas prices, conventional gas wells have all but disappeared.

**Figure 4.4:** Number of Spuds by Well Type (right axis) and Oil and Gas Prices (left axis), 2000-2015, Quarterly

*Sources:* Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

As shown in the figure, gas drilling fell dramatically after the collapse of gas

---

25 Oil prices represent West Texas Intermediate (WTI) prices, divided by 5.8 to convert to dollars per million British thermal units. Both oil and gas prices are the simple average of the next 12 months of futures prices in real 2014 dollars.
prices in 2008. Natural gas prices continued to decline in the subsequent years, during which the number of conventional gas wells drilled fell by 70 percent, while unconventional wells drilling fell by about 8 percent. We explore the reasons for this asymmetric relationship in detail in section 4.4.

To conclude, while the data focus on only Texas gas wells, these figures and summary statistics illustrate that our dataset is consistent with commonly-discussed trends in the industry and detailed conversations with industry participants.

4.4 Models and Results

4.4.1 Overview

We divide our analysis of the natural gas production process into three stages: the decision to commence drilling (or “spud”) a well, how quickly to complete and start production at a well (conditional on spudding), and how quickly to extract gas from the well (conditional on first producing a well). We describe our analysis of the spud decision in section 4.4.2, the completion and production start decision in section 4.4.3, and the profile of the quantity produced in section 4.4.4. Finally, we integrate the analysis of each of these different stages in section 4.4.5.

4.4.2 Stage 1: Commence Drilling (Spud) a Well

Drilling Estimation Method

Our empirical specification represents drilling activity as a log-linear approximation that depends on the determinants of the profits from drilling. These determinants include the expected present value of the future stream of revenues and costs associated with different types of gas wells. This reflects an upward-sloping supply curve, which could be due to an upward-sloping supply of rig rentals (as in Anderson et al. 26 These numbers are the changes in total wells drilled for each well type in 2012, relative to 2010.)
2014), heterogeneity in drilling costs, and/or heterogeneity in well productivity (as
in Metcalf 2017). Because gas wells may produce oil in addition to gas (see Table
4.1), oil prices may affect the incentive to drill gas wells. For this reason, we include
both gas prices and oil prices as potential factors affecting gas drilling investment.

We estimate the relationship in first differences because both drilling activity and
our oil and gas revenue variables are non-stationary time series, but are stationary
in first differences. This leads to the following specification for the number of gas
wells drilled:

$$\Delta \ln{p_t} = \beta_0 + \sum_{l=0}^{L} [\beta_{1,t} \Delta \ln(p_{gas,t-l} q_{gas,t-l}) + \beta_{2,t} \Delta \ln(p_{oil,t-l} q_{oil,t-l})] + \gamma' \Delta X_t + \varepsilon_t \quad (4.1)$$

where $w_t$ represents the number of wells that are spudded during period $t$, which
is measured in quarters. Expected revenue in each period is equal to a well’s pro-
ductivity (denoted $\bar{q}_{gas,t}$ and $\bar{q}_{oil,t}$ respectively) multiplied by the price ($\bar{p}_{gas,t}$ and
$\bar{p}_{oil,t}$). $X_t$ is a vector of controls for wells drilled in period $t$. This includes quarterly
indicators in all specifications to control for seasonality in drilling activity. In one
specification, it also includes our productivity variables. The primary parameters
of interest are $\beta_{1,t}$, which represent the $l$-lagged elasticity of gas well drilling. We
include $L = 2$ quarterly lags of the revenue variables to account for the fact that
drilling a well takes time due to the need to obtain mineral rights, drilling permits,
contracts with drilling service companies, and transport drilling rigs to well sites.
In this specification, the cumulative drilling response with respect to gas prices27 is
given by $\sum_{l=0}^{L} \beta_{1,t}$.

The advantage of using expected revenue as an explanatory variable, rather than
simply gas and oil prices as in many other studies, is twofold. First, it more closely
reflects the reality that the incentive to drill a well depends on the well’s total rev-

27 We do not distinguish between “revenue” or “price” responses since they are equivalent holding
productivity constant, as in the case of an exogenous price shock. This is due to the equality
$\beta_{1,t} \ln(p_{gas,t-l} q_{gas,t-l}) = \beta_{1,t} \ln(p_{gas,t-l}) + \beta_{1,t} \ln(q_{gas,t-l}).$ Intuitively, a 1 percent increase in prices
is equivalent to a 1 percent increase in revenues, holding well productivity constant. Nonetheless,
our sensitivity analysis shows that the estimated “price” response is very similar.
venue, which depends in turn on the well’s overall production rather than the price of one unit of its output. Second, the changes to well productivity contribute an additional source of revenue variation in the data. Intuitively, a 10 percent increase in production has the same revenue impact as a 10 percent increase in output prices.

For the expected future prices of gas and oil $\hat{p}$, we use the simple average of the next 12 months of futures prices for Henry Hub natural gas and WTI oil, adjusted for realized inflation (see section 4.3.2). This captures the fact that drilling decisions yield a stream of production (and profits) over the course of the coming months into the future. Because firms discount future revenues and well output declines rapidly over time (see Table 4.1), much of the present value of revenue from oil and gas wells depends on prices in the first year or so of production.28,29

Because we do not observe firms’ expectations about well productivity ($\hat{q}_{gas,t}$ and $\hat{q}_{oil,t}$), we proxy for it using a measure of initial production from recent gas wells, which we refer to as “productivity” for short. Specifically, we use the average initial production values (first full month) for wells drilled in the prior two quarters.30 Our preferred specifications do not include controls, but including average well depth and lateral lengths as cost controls has little effect on the results.31

We consider the issue of endogeneity of natural gas prices due to unobserved supply shocks (e.g., drilling cost shocks) that simultaneously increase drilling but

28 We do not explicitly apply any time discounting to our analysis because that would require making assumptions about firms’ discount rates. We simply argue that focusing on one year of prices is reasonable because revenues from the first year of production dominate the net present value of a well’s cash flow.

29 Using futures prices as a measure of price expectations may be problematic if futures prices differ from expected spot prices. Baumeister and Kilian (2016b) (and Baumeister and Kilian (2016a)) show that this difference can be substantial for long-dated oil futures contracts. The combination of discounting and approximately exponential decline curves somewhat mutes the impact of this distinction on the expected net present value of revenues, which helps to justify our approach.

30 This data is presented in the appendix. Because production profiles decline quickly (Table 4.1), production in the first full month is a fairly reliable indicator of a well’s long-term output. The correlation between first-month production and first-year production is 0.89, and for mature wells in the data that have produced the vast majority of their output, the correlation between first-month production and cumulative production is 0.72.

31 The cost measures chosen are intuitive and comport with suggestions from industry operators.
reduce prices, threatening to bias our elasticity estimates towards zero. The need to instrument for price endogeneity is widely recognized in the literature estimating demand elasticities for fossil fuels (Davis and Kilian 2011b and Coglianese et al. 2017), so similar issues may also arise when estimating supply elasticities.\footnote{Studies of U.S. oil supply elasticities often do not instrument for oil prices based on the historically plausible argument that incremental production from the United States (and Texas in particular) is small relative to the global oil market. This argument requires that price-altering cost shocks are external in nature to Texas firms, such as supply disruptions from the Middle East. However plausible this argument has been for oil, it is certainly less sound for natural gas, which is primarily a North American market and has been strongly affected by the shale gas boom.}

In particular, the shale revolution has arguably affected oil prices in recent years, raising concerns about the endogeneity of oil prices. Indeed, Kilian (2016), Kilian (2017a), and Kilian (2017b) find that the impact of the shale revolution was important for WTI prices during 2011 through mid-2014, although less important thereafter. For these reasons, we instrument for both of our oil and gas revenue variables (both contemporaneous and lagged).

We use four instruments: U.S. population-weighted heating degree days (HDD), cooling degree days (CDD), lagged U.S. working gas inventories, and the Commodity Research Bureau (CRB) Raw Industrial Commodity Index.\footnote{HDD, CDD, and gas inventory data is from EIA, available at \url{http://www.eia.gov/forecasts/steo/query/}. The CRB Index was collected from Bloomberg L.P. It is an index tracking the prices of the following raw industrial commodities: “burlap, copper scrap, cotton, hides, lead scrap, print cloth, rosin, rubber, steel scrap, tallow, tin, wool tops, and zinc” (source: CRB). We also considered other instruments, including offshore gas well shut-ins due to hurricane events, hydropower production, gas production from other regions, non-farm employment, and a gas-weighted industrial production index. While each of these has the potential to shift the demand for onshore gas production (or to act as a proxy thereof), we do not use them because proved to be weak instruments. Other instruments explored include U.S. urbanization rates, the number of homes with natural gas connections, and gas pipeline network density, all of which were not available at the quarterly time step required in our analysis.}

We include both contemporaneous values and up to four quarterly lags of these instruments (as suggested by a reviewer), with the exception of gas inventories for which we only use the lagged values because the contemporaneous value is likely endogenous to gas prices.

The first three instruments are standard gas demand shifters. HDD and CDD are commonly-used energy demand shifters. The use of lagged inventories relates conceptually to Hausman and Kellogg (2015) and Roberts and Schlenkera (2013),
although those studies use a different IV strategy. The estimated price response is robust to dropping the inventory instrument.

The final instrument, the CRB Raw Industrial Commodity Index is included as a proxy for global commodity demand. Baumeister and Kilian (2012) show that this index is a good predictor of oil prices. Barsky and Kilian (2001) previously noted that oil prices tend to move together with the prices of other industrial commodities, suggesting common demand factors underlie their co-movements. The CRB index instrument is important for the power of our first stage–dropping it results in substantially smaller first stage F statistics.

We also present estimation results separately for unconventional and conventional gas wells to test whether the drilling response is different for unconventional wells. Unless otherwise stated, the sample period is 2000 Q1 to 2015 Q3, and the unit of observation is one quarter. All equations are estimated using two stage least squares (2SLS) with Newey-West HAC robust standard errors.

**Drilling Estimation Results**

Table 4.2 shows the results of estimating equation (4.1). The first three rows report the estimated price elasticity of drilling for 0, 1, and 2 quarterly lags of gas revenues (or prices in some specifications). The sum of these effects is the cumulative price response after three quarters (and revenue response more generally), also reported in the bottom half of the table along with standard errors. Our preferred specification is the simplest one, found in column (1), finding a statistically significant \( p = 0.02 \) cumulative drilling response of 0.9. If instruments are used only for current (and not lagged) revenues, the standard errors shrink and the significance level increases considerably.

---

34 The CRB index replaces the instrument of copper prices used in a previous draft because the CRB index, suggested by a reviewer, proved to be a stronger instrument.

35 In a similar vein, Hamilton (2014) uses copper prices to proxy industrial commodity demand, including oil.

36 In this case, the standard errors shrink by more than half, leading to a z statistic of 4.6 and \( p < 0.001 \).
A concern is that revenues may be endogenous through well productivity. While our six-month-moving-average method of estimating productivity aims to avoid this problem, we present two sensitivities regarding well productivity.

First, in column (2) we estimate drilling activity as a simple function of prices (rather than revenues), which are not subject to the concern about the endogeneity of productivity. In this specification, the key parameter of interest—the cumulative gas price response—is a somewhat smaller 0.7 but is statistically indistinguishable from the previous estimate (0.9 in column (1)).

Second, in column (3) we allow productivity to enter the equation separately to allow for the possibility that drilling investment responds differently to changes in prices versus changes in productivity, as suggested by a reviewer. Here, the cumulative gas price response is estimated to be 0.8, again statistically indistinguishable from the base model in column (1). Note that the specification in column (1) is simply column (3) with the constraint that the price and productivity coefficients are equal. In that context, we cannot reject at the 5 percent level the hypothesis that the associated price and productivity coefficients are the same. This is true for tests of both the gas and oil parameters, either testing the lags individually or jointly. However, this failure to reject is in part due to the large standard errors on the productivity coefficients.

Based on these results, we prefer the specification in column (1) for several reasons. First, we cannot reject the equality of the price and productivity coefficients. Second, the revenue variables more accurately represent revenues received by firms, and hence better represent firm incentives. Third, changing productivity is an important factor explaining why unconventional drilling continues to be profitable despite falling natural gas prices. This means rising expected revenues per well is important in explaining recent fluctuations in drilling activity, particularly for unconventional wells.

Columns (4) and (5) show results from re-estimating equation (4.1) separately for unconventional and conventional gas wells, including computing well productivity.
and revenues separately for each of these well types.\textsuperscript{37} Despite using very different drilling activity and productivity variables, the estimated elasticities are similar to each other (about 0.7) with overlapping confidence intervals. From this we conclude that there is no obvious difference between the drilling elasticity for unconventional and conventional gas wells, supporting our preferred specification in column (1), which considers both well types together.

The Wu-Hausman tests reject the null of no endogeneity in all specifications. Moreover, Sargan tests cannot reject the null that our overidentifying restrictions are valid, with one exception (the conventional-only specification). This supports the need for and validity of the IV approach. As a sensitivity, we also estimated equation 4.1 using OLS, ignoring price endogeneity; consistent with theory suggesting a downward-biased estimate, we found a substantially smaller gas price response (0.5 versus 0.9).

Our estimate, 0.9, is similar to those derived from other recent sources. Hausman and Kellogg (2015) estimate a cumulative drilling response of 0.81 using a different methodology and smaller sample.\textsuperscript{38} Using data from Wyoming, Mason and Roberts (2016) estimate a gas drilling elasticity of 0.61 to 0.82, depending on the specification. In addition, the long-run gas supply elasticity implied by the EIA’s Annual Energy Outlook 2015 is approximately 0.5.

However, none of those analyses estimate separate price responses for unconventional and conventional natural gas, which we do.

In the next two sections we build on the analysis of the drilling decision assessed in this section and consider the subsequent stages of the production process. In section 4.4.3 we analyze the amount of time it takes for a well to begin producing

\textsuperscript{37} For column (4)–the unconventional-only estimation–we focus on the 2005-2015 sample period for the reasons described above. Further, the small number of unconventional wells in the early 2000s leads to noisy and unreliable productivity estimates for revenues before 2005. As suggested by a reviewer, we re-estimated the conventional equation in column (5) using the same time period (2005-2015), finding that the estimated conventional elasticity increases somewhat, but it is nonetheless well within the range of sampling uncertainty.

\textsuperscript{38} That paper estimated the relationship in levels, rather than differences. Among other differences between our analyses, they used a different data source that could not distinguish between unconventional and conventional drilling.
once it is drilled. In section 4.4.4 we analyze the time profile of wells’ gas production.

Table 4.2: Drilling Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta \log(\text{Gas Revenues}))</td>
<td>0.35</td>
<td>0.33</td>
<td>0.29</td>
<td>0.29</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.14)</td>
<td>(0.18)</td>
<td>(0.12)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>(\Delta \log(\text{Gas Revenues}), 1 \text{ Lag})</td>
<td>0.37</td>
<td>0.29</td>
<td>0.36</td>
<td>0.39</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.050)</td>
<td>(0.15)</td>
<td>(0.18)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>(\Delta \log(\text{Gas Revenues}), 2 \text{ Lags})</td>
<td>0.17</td>
<td>0.070</td>
<td>0.15</td>
<td>0.0023</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.062)</td>
<td>(0.081)</td>
<td>(0.10)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>(\Delta \log(\text{Oil Revenues}))</td>
<td>0.41</td>
<td>0.57</td>
<td>0.54</td>
<td>0.28</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.12)</td>
<td>(0.17)</td>
<td>(0.19)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>(\Delta \log(\text{Oil Revenues}), 1 \text{ Lag})</td>
<td>-0.089</td>
<td>0.13</td>
<td>0.11</td>
<td>0.081</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.093)</td>
<td>(0.12)</td>
<td>(0.13)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>(\Delta \log(\text{Oil Revenues}), 2 \text{ Lags})</td>
<td>0.25</td>
<td>0.46</td>
<td>0.41</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.095)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>(\Delta \log(\text{Gas Productivity}))</td>
<td>1.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta \log(\text{Gas Productivity}), 1 \text{ Lag})</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta \log(\text{Gas Productivity}), 2 \text{ Lags})</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta \log(\text{Oil Productivity}))</td>
<td>-0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta \log(\text{Oil Productivity}), 1 \text{ Lag})</td>
<td>-0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta \log(\text{Oil Productivity}), 2 \text{ Lags})</td>
<td>-0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Gas Price Response</td>
<td>0.89</td>
<td>0.70</td>
<td>0.80</td>
<td>0.68</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.13)</td>
<td>(0.27)</td>
<td>(0.33)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>Cumulative Oil Price Response</td>
<td>0.56</td>
<td>1.17</td>
<td>1.06</td>
<td>0.59</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.19)</td>
<td>(0.24)</td>
<td>(0.33)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Observations (Quarters)</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>43</td>
<td>63</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.32</td>
<td>0.49</td>
<td>0.32</td>
<td>0.30</td>
<td>0.47</td>
</tr>
<tr>
<td>Revenues or Prices Used?</td>
<td>Revenues</td>
<td>Prices</td>
<td>Prices</td>
<td>Revenues</td>
<td>Revenues</td>
</tr>
<tr>
<td>Wells Included</td>
<td>All Gas</td>
<td>All Gas</td>
<td>All Gas</td>
<td>Unconv. Gas</td>
<td>Conv. Gas</td>
</tr>
<tr>
<td>First-stage F Statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas Revenues/Prices</td>
<td>30.4</td>
<td>11.2</td>
<td>11.2</td>
<td>23.2</td>
<td>31.8</td>
</tr>
<tr>
<td></td>
<td>31.1</td>
<td>30.7</td>
<td>30.7</td>
<td>152.3</td>
<td>52.2</td>
</tr>
<tr>
<td>Gas Revenues/Prices, 2 Lags</td>
<td>41.1</td>
<td>39.8</td>
<td>39.8</td>
<td>191.1</td>
<td>28.4</td>
</tr>
<tr>
<td>Oil Revenues/Prices</td>
<td>29.6</td>
<td>46.4</td>
<td>46.4</td>
<td>22.2</td>
<td>19.9</td>
</tr>
<tr>
<td>Oil Revenues/Prices, 1 Lag</td>
<td>40.2</td>
<td>39.6</td>
<td>39.6</td>
<td>31.2</td>
<td>16.8</td>
</tr>
<tr>
<td>Oil Revenues/Prices, 2 Lags</td>
<td>111.0</td>
<td>210.4</td>
<td>210.4</td>
<td>43.1</td>
<td>62.3</td>
</tr>
<tr>
<td>p-value, Wu-Hausman test for endogeneity</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.005</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>p-value, Sargan overidentification test</td>
<td>0.32</td>
<td>0.42</td>
<td>0.31</td>
<td>0.30</td>
<td>0.019</td>
</tr>
</tbody>
</table>

Notes: Newey-West HAC standard errors in parentheses. All specifications include quarterly indicators, which are suppressed for space. Sample period is 2000-2015, quarterly, with the exception of (5) which is estimated over 2005-2015 because that is when the shale gas boom began in earnest. Instruments are Cooling Degree Days (CDD), Heating Degree Days (HDD), copper prices, and gas inventories. With the exception of gas inventories, contemporaneous and 1-4 lagged values are used as instruments. Contemporaneous gas inventories are not used, as these are likely endogenous to gas prices. In (5) and (6), the computation of revenues is performed separately for unconventional and conventional wells to account for differences in productivity trends between them.
4.4.3 Stage 2: Spud-to-Production Time

Duration Model

Once a well is spudded, production does not typically begin until at least a month later (see Table 4.1), and we find that there is considerable variation in how long each well takes to reach initial production. After spudding the well, the time to finish drilling the well depends on factors like well depth and the length of well laterals. The completion stage follows the drilling stage and may require artificially fracturing the well, which also increases the time to production. After completion, the well can begin producing natural gas and oil. Other factors that might influence completion times are fuel prices—whereby operators work faster when prices are higher to achieve revenues earlier—and logistics like renting a completion rig. As previously mentioned in section 4.3.3, unconventional wells in our data tend to take longer to reach production, due to the additional labor required to horizontally drill and fracture.

We estimate the distribution of spud-to-completion times as a function of prices using duration models (i.e., survival time or hazard models) with time-varying coefficients, as described in the online appendix. We estimate this model using maximum likelihood and assuming a generalized gamma distribution for the underlying density of the baseline hazard function, which is parameterized by two ancillary parameters determining the shape of the distribution.

A hazard model estimated with the gamma distribution is parameterized in an “accelerated failure time” (AFT) setup. In this setup, the explanatory variables can be interpreted as additively affecting the observation’s logged expected “failure” time, which here is the time it takes for a spudded well to reach production. This implies that if an increase in gas prices encourages operators to work faster to speed up the time of production, this would be represented as a negative coefficient in the AFT model (i.e., reducing the time to production).

We estimate the hazard models for each unconventional and conventional wells independently, with no cross-equation restrictions. This allows each well type to have
its own distribution, in addition to its own estimated parameters. As in the drilling regressions in section 4.4.2, the explanatory variables include (in logs): expected revenue from gas and oil production,\textsuperscript{39} and cost controls of well depths and lateral lengths.\textsuperscript{40} As a robustness check, we also present specifications using simply oil and gas prices instead of revenues.

The unit of observation is a well-month, which means we can use well-specific characteristics (lateral length and well depth) rather than the means across wells used in the drilling equations of the prior section. Nonetheless, even though we observe wells’ realized production, we do not use it because firms do not know with certainty how productive a well will be until it actually starts producing. Hence, using the well’s actual, \textit{ex post} production as an explanatory variable would imply that the firm responds to unobservable information. Instead, we use the same method of calculating productivity and revenues used in the drilling equations.\textsuperscript{41} In the terminology of hazard analysis, we consider a well to be “at risk” of being produced for 24 months following its spud month, at which point it exits our sample.\textsuperscript{42}

**Spud-to-Production Duration Estimation Results**

Table 4.3 shows the estimates of the duration models. Our preferred estimates are again the simplest ones, shown in columns (1) and (5) for unconventional and conventional wells respectively. The first coefficient reported in column (1) is -0.23, indicating that a 10 percent increase in the gas price (or revenues more generally) is expected to reduce the time to production for an unconventional well by about 2

\textsuperscript{39} As before, the results are robust to simply using gas and oil prices, but we use the revenue variables as we believe they are more accurate depictions of firms’ incentives. We do not use instrumental variables as the theoretical and empirical literature for implementing them in duration models is very limited. We are aware of only one published paper, MacKenzie et al. (2014), that considers the implementation of instrumental variables in duration models; that study only applies to a Cox proportional hazard setup, as opposed to the accelerated failure time model that we use.

\textsuperscript{40} All specifications also have spud year fixed effects, controlling for secular trends in drilling and completion times.

\textsuperscript{41} Since we conduct the hazard analysis at the monthly level, average revenues are calculated on a monthly basis, rather than quarterly.

\textsuperscript{42} We chose 24 months for the reasons described in section 4.3.2
percent. That only amounts to about 3 days of the 4.7 month average completion time for unconventional wells (see Table 4.1). Across columns (1) through (4), unconventional wells have generally consistent price response coefficients between -0.19 and -0.27. The corresponding responses for conventional wells, found in columns (5) through (8), are somewhat more modest, with coefficients in the range of -0.02 to -0.08.43

Table 4.3: Spud-to-Production Duration Model Results

<table>
<thead>
<tr>
<th>Spud-to-Production Survival Time</th>
<th>Unconventional Gas Wells</th>
<th>Conventional Gas Wells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(Gas Revenues or Prices)</td>
<td>-0.228</td>
<td>-0.083</td>
</tr>
<tr>
<td></td>
<td>(0.0205)</td>
<td>(0.0208)</td>
</tr>
<tr>
<td>Log(Oil Revenues or Prices)</td>
<td>-0.194</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>(0.0262)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Log(Gas Productivity)</td>
<td>0.160</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>(0.0147)</td>
<td>(0.0222)</td>
</tr>
<tr>
<td>Log(Oil Productivity)</td>
<td>0.059</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>(0.0218)</td>
<td>(0.0242)</td>
</tr>
<tr>
<td>Log(Vertical Depth)</td>
<td>0.174</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.0363)</td>
<td>(0.0610)</td>
</tr>
<tr>
<td>Log(Lateral Length)</td>
<td>-0.021</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>(0.0161)</td>
<td>(0.00751)</td>
</tr>
<tr>
<td>Gamma Density Function Parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revenues or Prices?</td>
<td>0.473</td>
<td>0.505</td>
</tr>
<tr>
<td></td>
<td>0.475</td>
<td>0.505</td>
</tr>
<tr>
<td></td>
<td>0.491</td>
<td>0.491</td>
</tr>
<tr>
<td></td>
<td>0.455</td>
<td>0.488</td>
</tr>
<tr>
<td>Revenues (N*T)</td>
<td>146,859</td>
<td>126,990</td>
</tr>
<tr>
<td></td>
<td>146,859</td>
<td>126,990</td>
</tr>
<tr>
<td></td>
<td>146,859</td>
<td>126,990</td>
</tr>
<tr>
<td></td>
<td>146,736</td>
<td>126,990</td>
</tr>
<tr>
<td>Wells (N)</td>
<td>25,725</td>
<td>36,055</td>
</tr>
<tr>
<td></td>
<td>25,725</td>
<td>36,055</td>
</tr>
<tr>
<td></td>
<td>25,725</td>
<td>36,055</td>
</tr>
<tr>
<td></td>
<td>25,701</td>
<td>36,055</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-18,744</td>
<td>-29,053</td>
</tr>
<tr>
<td></td>
<td>-18,767</td>
<td>-29,073</td>
</tr>
<tr>
<td></td>
<td>-18,717</td>
<td>-29,053</td>
</tr>
<tr>
<td></td>
<td>-17,890</td>
<td>-28,498</td>
</tr>
</tbody>
</table>
| p-value on test of equal unconv./conv. gas elasticities | <0.0001 | <0.0001 | <0.0001 | <0.0001

43 These results are robust to including lagged oil and gas revenues. The coefficients on the lagged values were small and generally insignificant. The sum of the contemporaneous and lagged coefficients were very similar to the coefficients reported in Table 4.3.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourage conventional natural gas efforts, potentially representing a substitution effect, while having more muted and ambiguous effects on unconventional effort. Columns (2)-(3) and (6)-(7) show that our estimated elasticities are robust to using simply oil and gas prices (rather than our measure of revenues) and to allowing prices and productivity to enter separately in the specification.

Sources: Authors’ calculations based on data from Drillinginfo, EIA, and Bloomberg

Increases in oil revenues seem to slightly discourag
Columns (4) and (8) add lateral lengths and well depths as controls. The coefficient on lateral lengths sensibly indicates that unconventional wells with longer laterals take longer to reach production, consistent with longer drilling and completion periods. Well depth plays an analogous role for conventional wells: deeper wells take longer to reach production.

While the gas price response coefficients are precisely estimated and their signs are consistent with economic expectations, they are small. For example, in order to achieve a 10 percent reduction in spud-to-production time for conventional wells, gas prices would have to nearly triple. For the more-responsive unconventional wells, gas prices would have to increase by more than 40 percent.\footnote{These required price changes represent the change in gas prices needed to shift the fitted distributions in the benchmark specifications, computed at covariate means, such that the mean of the distribution decreases by 10 percent.}

This lack of strong price response is illustrated in Figure 4.5. That figure plots the parameterized gamma distributions for each well type under alternative natural gas price assumptions of $3.00 and $6.00 per million Btu.\footnote{This figure assumes 2015 average productivity values and a $50 per barrel oil price.} The effects are not large. Despite an assumed doubling of natural gas prices, the effect on the spud-to-production time distribution is modest.

These results suggest that once drilling has commenced, gas prices do not strongly influence the decision about whether and how fast to begin producing a well. This is sensible, since once drilling has begun, much of the well development costs have been sunk. There also may be limited opportunities for speeding up the completion process beyond a certain point.\footnote{Although there is always the option to slow down or stop the completion process in the face of low prices, this may not save costs if service contracts are already in place.} With low additional marginal costs of production, operators typically have strong incentives to begin production as soon as possible.
4.4.4 Stage 3: Production Profile over Time

Production Profile Estimation Method

The final stage of the gas production process is the flow of gas from wells once they begin producing, and how that flow evolves over time. In this section, we estimate the time profile of well-level gas production and its relationship with prices.

Once a well begins producing, it often produces gas and oil for many years, with the production profile being determined principally by reservoir pressure. Because the variable cost of production from an existing well is very low, an operator would typically want to produce oil and gas at a well’s capacity. For this reason, we would not expect gas and oil prices to have a significant effect on production from existing wells by a competitive firm. Instead, a well’s flow rate is largely determined by the
amount of pressure left in the reservoir to force the resource to the surface. The flow rate is therefore generally out of the operator’s control.\textsuperscript{47} These arguments are analyzed at an aggregate level for oil in Anderson et al. (2014), and we find similar results for gas at the well level.

Even if production from existing wells is not price responsive itself, understanding the time profile of production is nonetheless still important to understanding aggregate supply responsiveness. This is because these profiles determine the relationship between drilling effort and realized production over time, and production profiles are quite different for unconventional versus conventional wells.

As described further in the next section, our estimates show a lack of price response of output from existing gas wells using a detailed panel dataset describing monthly gas production for each well’s productive life. Specifically, we run fixed-effects regressions of the form:

\[
\ln(q_{i,\text{gas},j,t}) = \chi_i + \eta_{\text{gas},j} \ln(p_{\text{gas},t}) + \eta_{\text{oil},j} \ln(p_{\text{oil},t}) + g_j(Age_{i,t}) + \varepsilon_{i,j,t},
\]

where \(i\) indexes the well and \(t\) indexes the calendar time in months. \(q_{i,\text{gas},j,t}\) is the gas production from well \(i\) of type \(j\) in month \(t\). \(\chi_i\) is a well-level fixed effect, which can roughly be interpreted as initial (log) production for well \(i\). \(p_{\text{gas},t}\) and \(p_{\text{oil},t}\) are prompt-month gas and oil prices. The parameters of interest are \(\eta_{\text{gas},j}\) describing the contemporaneous price elasticity of gas production from a well (of type \(j\)) that has already been drilled. We use spot (i.e., prompt-month) oil and natural gas prices because of the immediate nature of the potential price response from existing wells.\textsuperscript{48} Given the well-level fixed effects, our identification of the price response comes from changes in prices during the life of a well. The discussion above and prior evidence suggests that this parameter would be estimated as being close to zero. Regardless, the different production profiles of wells can still be consequential for the overall supply responsiveness to prices because this stage is conditioned on

\textsuperscript{47} There are some exceptions. Firms can extract more hydrocarbons through investments in enhanced recovery methods like pumps (for oil) and various injection methods. For unconventional reserves, firms have the added option to re-fracture the well, which can create new fissures that release more hydrocarbons to the surface.

\textsuperscript{48} Using lagged prices does not substantially affect the results.
the earlier drilling decision, which we found above is responsive to price.

$Age_{i,t}$ is the age of well $i$ at time $t$ (i.e., the number of months since it began production). $g_{j}(Age_{i,t})$ is a function of the age of a well of type $j$, and we allow for flexible production profiles by approximating this function using polynomials of varying degrees as well as a cubic spline.\(^49\) This flexible trend controls for time series concerns.\(^50\) We drop the first month of production, as wells are often only operating for a fraction of this month, instead beginning with the first full month of production. The age function is indexed by $j$ to allow the average production profile to vary based on well type (i.e., unconventional versus conventional). Standard errors are clustered at the well level.

**Production Profile Estimation Results**

Table 4.4 contains the results for the fixed effects regressions for unconventional and conventional gas wells. Consistent with the above discussion, we find very small coefficients on natural gas prices, suggesting that gas production from existing wells is not price responsive. The elasticity point estimates are small, typically ranging between $+0.07$ and $-0.03$ (with one exception in column (5), discussed below), all very close to zero and often of a theoretically implausible sign. The same is true for oil price coefficients, which are generally very close to zero and occasionally insignificant despite small standard errors.\(^51\)

The substantial size of our dataset (over 5 million well-month observations for conventional and unconventional wells combined) generates very small standard errors. As a result, even many of these very small estimates (e.g., smaller than 0.02 in magnitude) are nevertheless significant at the 1 percent level. These negligible

\(^49\) The cubic spline uses knots at every 12-month interval after initial production.

\(^50\) We conducted statistical tests for non-stationarity in the well-production panel, finding strong rejection of a panel unit root (using Stata’s `xtunitroot` command). Nonetheless, we also estimated these specifications in first differences, finding qualitatively similar results with even smaller estimated price elasticities. We present the results from estimation in levels to ease the reader’s interpretation of the decline curve coefficients.

\(^51\) Excluding oil prices from the specification completely (not shown) also has little effect on the results.
Table 4.4: Well Production Profile Fixed Effects Regressions

<table>
<thead>
<tr>
<th>Dep. Var.: Log(Gas Production)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconventional Wells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Gas Price)</td>
<td>0.07</td>
<td>-0.01</td>
<td>-0.03</td>
<td>-0.02</td>
<td>0.07</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Log(Oil Price)</td>
<td>-0.07</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Well Age (months)</td>
<td>-0.021</td>
<td>-0.044</td>
<td>-0.067</td>
<td>-0.086</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000131)</td>
<td>(0.000274)</td>
<td>(0.000475)</td>
<td>(0.000749)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Age^2 (months)</td>
<td>0.00023</td>
<td>0.00076</td>
<td>0.00153</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000003)</td>
<td>(0.00001)</td>
<td>(0.000028)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Age^3 (months)</td>
<td>-3.2E-06</td>
<td>-1.4E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.1E-08</td>
<td>(3.7E-07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Age^4 (months)</td>
<td></td>
<td></td>
<td></td>
<td>4.0E-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.6E-09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Well Age)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.679</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00301)</td>
<td></td>
</tr>
<tr>
<td>Conventional Wells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Gas Price)</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.03</td>
<td>0.20</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Log(Oil Price)</td>
<td>-0.18</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.06</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Well Age (months)</td>
<td>-0.017</td>
<td>-0.035</td>
<td>-0.052</td>
<td>-0.069</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000077)</td>
<td>(0.000168)</td>
<td>(0.000283)</td>
<td>(0.000433)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Age^2 (months)</td>
<td>0.00011</td>
<td>0.00039</td>
<td>0.00085</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000001)</td>
<td>(0.000004)</td>
<td>(0.000011)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Age^3 (months)</td>
<td>-1.1E-06</td>
<td>-5.4E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.6E-08</td>
<td>(9.7E-08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Age^4 (months)</td>
<td>1.0E-08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.0E-10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Well Age)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.761</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00241)</td>
<td></td>
</tr>
<tr>
<td>N (Well-Months)</td>
<td>5,331,586</td>
<td>5,331,586</td>
<td>5,331,586</td>
<td>5,331,586</td>
<td>5,331,586</td>
<td>5,331,586</td>
</tr>
<tr>
<td>Number of Wells</td>
<td>62,568</td>
<td>62,568</td>
<td>62,568</td>
<td>62,568</td>
<td>62,568</td>
<td>62,568</td>
</tr>
<tr>
<td>Well Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cubic Spline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared (Full Model)</td>
<td>0.744</td>
<td>0.759</td>
<td>0.762</td>
<td>0.763</td>
<td>0.762</td>
<td>0.764</td>
</tr>
<tr>
<td>R-Squared (Excluding Fixed Effects)</td>
<td>0.352</td>
<td>0.388</td>
<td>0.397</td>
<td>0.400</td>
<td>0.396</td>
<td>0.401</td>
</tr>
</tbody>
</table>

Standard errors in parentheses, clustered at the well level. The first month of each well's production is dropped, operational for only a fraction of its first month.

Sources: Authors' calculations based on data from Drillinginfo, EIA, and Bloomberg

elasticity estimates are generally robust to variations in the functional form describing the decline path. Polynomials of orders one through four and cubic splines produce similar results. The linear-in-age specification in column (1) is probably
insufficiently flexible because production is commonly observed to decline slower-than-exponentially (as captured by a positive “b” parameter in the Arps equation), and a linear specification effectively assumes an exponential decline. We find support for slower-than-exponential declines in the significant positive coefficient on the well-age-squared terms in columns (2) through (4). We also include a specification using log\( \log(Age_{i,t}) \), following Patzek et al. (2013) who argue for such a specification for unconventional gas wells in the Barnett shale formation, a subset of our data.

In general, we do not find a positive price response of gas production from existing wells. We also tested whether this price response may vary over the life of the well by estimating separate elasticities for every three-month bin (0-3 months, 4-6 months, etc.). These estimated elasticities generally varied in sign and were close to zero (e.g., smaller than 0.1) and were generally statistically insignificant, despite relatively small standard errors.

While Anderson et al. (2014) present time-series evidence showing that aggregate Texas oil production from existing wells is not price responsive, to our knowledge our study is the first to provide evidence for this proposition for natural gas production using well-level data. Because the negligible observed price response from existing wells comports with both structural economic factors and empirical evidence, we proceed to model production from existing wells as completely unresponsive to price. This assumption also allows us to model well production profiles non-parametrically in our combined model in section 4.4.5.

To represent the production profile from unconventional and conventional wells, we use the mean monthly production in our dataset by well age. For example, to estimate the production from an average unconventional well in its 7\(^{th}\) month of production.

\[^{52}\text{Patzek et al. (2013) specifically argue that the coefficient on log}(Age_{i,t})\text{ should be approximately -0.50 for unconventional gas wells, but we find a somewhat larger coefficient of -0.68. If we include each wells’ first, partial month of production close to -0.50. However, including the first, partial month is inappropriate when fitting parametric decline curves, as an inspection of Figure 4.6 suggests. Regardless, our finding of no meaningful price response is generally robust to including wells’ first, partial months. The only specification with an economically meaningful gas price response estimate is the one using log}(Age_{i,t})\text{ for conventional reservoirs, which is likely inappropriate because Patzek et al. (2013) only advise using that functional form for unconventional Barnett wells, not conventional ones.}\]

124
production, we calculate the average production from every unconventional well in our dataset during its 7th month. This is practically equivalent to estimating a regression of gas production on monthly indicators for well age. The results of this procedure are illustrated in Figure 4.6. The solid lines in that figure represent the mean production profile of unconventional (darker) and conventional (lighter) wells. The dashed lines represent the median production profiles.

**Figure 4.6:** Mean and Median Profile of Monthly Gas Production from Gas Wells

*Sources:* Authors’ calculations based on data from Drillinginfo and EIA

These production profiles reflect the well characteristics presented in Table 4.1. The mean and median unconventional production profiles strongly resemble each other, suggesting that unconventional gas production profiles are generally not heavily skewed in productivity. In contrast, the median conventional well produces much

53 For each well that has ceased production before the end of our sample period, we impute zero production in each month after its last month of production when calculating this average.
less than the mean, indicating a right-skewed distribution with relatively more “dry holes.” These facts align with discussions of decline paths in industry and popular press.

A concern is that these estimated production profiles suffer from compositional effects, with more recent wells have different productivity levels and also have a shorter history of data. We would not expect this effect to be strong for conventional wells because those wells have not exhibited strong trends in their productivity. Rising unconventional productivity makes this issue more of a concern. We consider this in the appendix by normalizing each well’s production by its peak production before computing average decline curves. This normalizes away changing productivity levels, eliminating the concern about compositional effects due to rising productivity. The resulting decline curves have similar shapes, indicating that the overall shapes of the curves are not being driven by compositional effects.

4.4.5 Integrating Natural Gas Supply Stages to Measure Overall Price Responsiveness

Integrated Model of Well Drilling, Production, and Decline over Time

In this section, we combine the analysis of the three gas supply stages from the preceding sections into an integrated simulation model. The purpose of this integrated model is to understand the overall price-responsiveness of natural gas supply, how it evolves over time, and how it differs between unconventional and conventional resources.

The three separate stages of the natural gas production process are the spud decision, the time from spudding to the first production of a well, and the production profile of a gas well over time. We link these stages together by simulating the effects of an unexpected, permanent percent shock to natural gas prices. We then illustrate the effect of this shock on gas drilling activity and natural gas production over time.
in a manner readily understandable in percentage terms.\footnote{\textsuperscript{54,55}}

This shock increases the number of spuds of each well type in every period, based on the model specified in equation (4.1) and the corresponding estimated price responses presented in Table 4.2. While this model is not a fully-specified dynamic model wherein producers intertemporally optimize extraction decisions, Metcalf (2017) shows how a similar model can be interpreted as an approximation to the case where the Hotelling-style reserve constraints are generally non-binding (e.g., if new reserve discoveries are continually added to the stock of potential wells over time). As Metcalf (2017) argues, the shale boom is exactly an example of that case.

The additional spudded gas wells take time to reach production, according to the estimated distributions associated with the hazard estimation results in Table 4.3 (illustrated in Figure 4.5). We then use the mean production profile shapes presented in Figure 4.6 to simulate the amount of production from wells of each type.

While in principle well productivity itself may respond to a price change, the direction of this effect is ambiguous and measuring it is beyond the scope of this study. On the one hand, firms may increase fracturing inputs as prices rise, resulting in higher average per-well production. Mason and Roberts (2016) find some evidence for a positive response from gas wells in Wyoming, although this result disappears when they control for operator, drilling direction, and reservoir characteristics. On the other hand, at higher prices, less productive wells may become profitable and are drilled, reducing average per-well production. Metcalf (2017) finds evidence for such a negative effect for oil drilling, but no effect for gas.

\footnote{This simulation focuses on gas supply from gas wells. Of course, gas is also produced as co-production from oil wells. This other component of supply is also important, amounting to as much as 30 percent of Texas gas production in 2015, but exploring it is beyond the scope of this paper. Other ongoing work has found that oil drilling is not responsive to gas prices however. A lack of responsiveness from co-production from oil-directed drilling would imply a total gas production response that is somewhat smaller than the 0.9 estimate found here for gas-directed drilling.}

\footnote{While we do not explicitly simulate the response to a price decrease, in our model such a response would be approximately symmetric to that of the price increase that we simulate. (Such a response is exactly symmetric for marginal price changes, but only approximately so for large changes due to the non-linearity of the exponentiation applied to the elasticity.) To approximate the response to such a decrease, one can multiply the incremental change in drilling or production by $-1$.}

127
A simple regression of quarterly changes in (log) gas productivity (data shown in the appendix) on quarterly changes in (log) gas prices (shown in Figure 4.4) yields a small and statistically insignificant elasticity of 0.038 (standard error: 0.042), suggesting little evidence for a strong net response of productivity to price changes, at least in the short run, consistent with the findings in Metcalf (2017). An anonymous reviewer suggested that this insignificant response may clarify the reasons underlying an upward-sloping drilling supply curve. If average well productivity does not strongly decline as prices rise, then it is not simply that unproductive wells are only drilled at higher prices, but rather that the marginal cost of drilling rigs is upward-sloping. Indeed, rental rates for drilling rigs are known to rise as drilling activity increases.

To conduct our simulation, we begin by noting that, given the number of gas wells drilled, the number of gas wells beginning production at month $t$ is the accumulation of the wells that were spudded in recent months. We can write this relationship precisely using the spud-to-production distributions in Table 4.3 (seen in equation (C.1) and illustrated in Figure 4.5). Denoting the discrete analogues of these distributions as $f_{j,l}$ for a well type $j$ beginning production $l$ months after spudding, we can write the number of wells beginning production at month $t$, denoted $x_{j,t}$, as a function of $f_{j,l}$ and the number of wells spudded in each of preceding 24 months, denoted $w_{j,t}$, as follows:

$$x_{j,t} = \sum_{l=0}^{24} w_{j,t-l} f_{j,l}. \quad (4.3)$$

Equation (4.3) thus combines supply stages 1 and 2: well drilling and commencement of production. Next, by combining this with stage 3—well-level production profiles—we can calculate the total gas production over time. As in the production profile analysis (see equation (4.2)), we use $q_{\text{gas},j,\tau}$ to denote gas production of a well of type $j$ in its $\tau$th month of operation. Denoting the productive life of a well of type $j$ as $T_j$, we can write total gas production at time $t$ from all wells of type $j$ as:
\[ Q_{gas,j,t} = \sum_{\tau=0}^{T_j-1} x_{j,(t-\tau)} q_{gas,j,\tau}. \]  

(4.4)

Using equations (4.3) and (4.4) and the results from the drilling, completion, and production models (sections 4.4.2, 4.4.3, and 4.4.4), we can simulate the effect of a change in prices on gas spuds \((w_{j,t})\), gas wells entering production over time \((x_{j,t})\), and their gas production over time \((Q_{gas,j,t})\).

A significant period of time is needed to reach a long-run equilibrium in gas production after a price shock because today’s production depends on events that happened as long at \(T_j\) periods ago. For example, if a typical well produces for 10 years, then today’s production level depends to some degree on drilling 10 years ago. This inertia is endemic to oil and gas supply dynamics, and it underpins much of the cyclicality in these markets. However, since old wells produce little, the effect of a price change is front-loaded. A key issue of interest for this paper is whether unconventional resources and technologies may reduce the volatility inherent in this industry characterized by large fixed investments and low variable production costs.

In our simulation’s baseline, we assume constant prices, implying that the quantity of drilling \(w_{j,t}\) does not vary over time, denoted \(w_{j}^{base}\). In each simulation, we specify \(w_{j}^{base}\) using the average observed values in the data (described in more detail below). Given this and equation (4.3), the number of wells entering production each month in the baseline equilibrium (i.e., without a price shock) is also constant:

\[ x_{j}^{base} = w_{j}^{base} \sum_{l=0}^{24} f_{j,l} = w_{j}^{base}, \]  

(4.5)

where the latter equality follows because the density of spud-to-production time must sum to one: \(\sum_{l=0}^{24} f_{j,l} = 1\).\(^{56}\) This and equation (4.4), in turn, imply that baseline gas production is also in long-run equilibrium in the baseline:

---

\(^{56}\) In reality, not all spudded wells go on to produce. However, for simplicity we are ignoring this complication. An exploration of Drillinginfo’s permit dataset found only a small fraction of the wells in our dataset for which we could identify as unconventional or conventional that were drilled but never reached production. For this reason, we ignore such wells. However, including them would simply multiplicatively scale down the price responsiveness for both well types slightly.
Using equations (4.3) through (4.6), we compute the number of gas wells beginning production and quantities of gas produced over time, by well type, under a baseline price scenario as well as under a scenario with a permanent price increase.

We conduct two sets of simulations. The first compares unconventional and conventional gas on a per-well basis, without accounting for the shift away from conventional drilling that has occurred during the shale revolution. The second simulation accounts for these changes over time, as well as changes in productivity, the level drilling, and the composition of well types. This latter simulation provides a more complete accounting of the effects of the shale revolution.

Simulation 1: Overall Unconventional vs. Conventional Gas Supply Responsiveness on a Per-well Basis

The first simulation shows the impacts of an unexpected, permanent 10 percent shock to natural gas prices (from $3.00 to $3.30 per million Btu). It is calibrated to 2015 baseline values of drilling activity, normalized to compare unconventional and conventional wells on a per-well basis. This ignores the shift away from conventional drilling in recent years, but allows for comparing the two types of resources on a per-well basis.

To do this, we run the simulation separately for unconventional and conventional wells, each time assuming the same spud baseline of 72 wells for each type: \( w_{\text{base conv}} = 72 \) or \( w_{\text{base unconv}} = 72 \).\(^{57}\) We use the same baseline for each well type to compare them on a per-well basis. This is intended to remove the crowding out of conventional wells by unconventional drilling in recent years and compare the two types of gas on an even footing.\(^{58}\) From these baselines, we simulate the time series of the number of

\[ Q_{\text{gas,j}}^{\text{base}} = x_j^{\text{base}} \sum_{\tau=0}^{T_j-1} q_{\text{gas,j,}\tau}. \]  

\(^{57}\) 72 is the average number of monthly gas spuds in our data during our final sample year, January 2015 - September 2015. The breakdown was 61 unconventional spuds and 11 conventional spuds.

\(^{58}\) The simulations in the next section account for the changing number of wells of each type drilled over time.
new gas wells entering production.

We convert these wells to aggregate gas production for each type of well using the estimated production profiles. We scale up the unconventional profile by a factor of approximately 1.2 to reflect the higher initial productivity levels of about 80,000 mcf per month in recent years (2010 to 2014), relative to average levels of about 67,000 over the entire 2005 to 2015 period.\textsuperscript{59} The result is a time series of changes in production by well type.

We then convert the change in gas wells beginning production and gas produced to percentage changes. To compare the results for unconventional and conventional wells on an equal footing, we compute percentage changes in gas production by dividing by the same denominator in each case: the amount of gas consistent with the types of gas wells actually drilled in our data, on average, in 2015.\textsuperscript{60} This equal denominator allows one to observe that, even if more conventional wells had been drilled in 2015 than actually were drilled (72 instead of 11), their much lower productivity means that they would still contribute relatively little to the overall gas supply response.

The simulation results are shown in Figure 4.7. The left panel shows the change in number of wells beginning production each month for unconventional and conventional wells as a percentage of the baseline number of wells. The price shock at period zero leads to new drilling effort which gradually bears fruit over the course of the subsequent 24 months and beyond. After 24 months, the wells reach a new long-run equilibrium, with the same number of wells beginning production in every month.

\textsuperscript{59} We exclude 2015 when computing recent productivity (of 80,000 mcf) out of concern that the large jump in productivity during that year reflects not permanent innovation but instead a temporary re-focusing of efforts on “sweet spots” during a time of low oil and gas prices.

\textsuperscript{60} The baseline gas production corresponds to the production from 61 unconventional gas wells and 11 conventional gas wells, computed in long-run equilibrium using the average production profiles used in this simulation and depicted in Figure 4.6. The amount is simply the cumulative total production across the average well’s lifetime, by well type, multiplied by the baseline number of gas spuds (61 and 11 for unconventional and conventional respectively). Note that our denominator includes only gas production from gas wells, and does not include gas co-production from oil wells. Hence, our estimates represent the response of gas production from gas-directed drilling. As of 2015, about 70 percent of total Texas gas production came from gas wells.
The right panel combines the results of the left panel with production decline paths and traces out the effect of the price shock on incremental natural gas supply, which shows a similar gradual adjustment. Each incremental well in the left panel produces for many years; as a result, the rising drilling effort builds on itself, until we reach a new long-run equilibrium after more than a decade. Before reaching the equilibrium, some portion of the production is from “legacy” wells drilled before the price shock. In other words, an immediate 9 percent increase in drilling effort increases total gas production by less than 9 percent because of the lack of response from legacy wells, at least until those wells eventually cease production and a new long-run equilibrium is reached. Only once the increased drilling effort has propagated throughout the system does it fully affect the level of gas produced.

Altogether, unconventional gas supply is in fact much more responsive to price changes than is conventional gas supply once one takes an integrated view of the entire production process. The time path to reach the new equilibrium depends on both the shape of the spud-to-production distributions and the production profiles. The more front-loaded the distribution and production profile are, the faster the drilling effort translates into production. Note, however, that because of the negligible price responsiveness of spud-to-production time and production from existing wells, the production response to the price shock is essentially entirely due to the

---

**Figure 4.7:** Change in Gas Wells Beginning Production and Natural Gas Produced from Those Wells, Following a 10 Percent Price Shock

*Sources:* Authors’ calculations based on data from Drillinginfo, Bloomberg, and EIA
effect on drilling investment.

These effects are partly, but not completely, offsetting for unconventional wells. On one hand, unconventional wells are more productive than conventional wells (see Figure 4.6). This supports the notion that unconventional gas production should respond more to a price change than conventional wells would. On the other hand, unconventional wells generally take longer to reach production (see Figure 4.5), which somewhat moderates the short-term effects of increased drilling effort on gas production. The superior productivity of unconventional wells more than offsets their longer drilling and completion times after just a few months. On net, for the first month of the simulation, unconventional gas actually responds less than conventional gas as wells take longer to bring online. But unconventional production quickly surpasses conventional production, once the much-more productive wells come online. By the 6th month of the simulation, the unconventional gas supply response is more than twice as large as the conventional response.

The right-hand panel of Figure 4.7 illustrates that, in the long run, the gas supply response from unconventional wells is about 2.7 times on a per-well basis ($\approx \frac{9.5 \text{ percent}}{3.5 \text{ percent}}$) larger than that of conventional wells. This is entirely due to the fact that unconventional wells are about 2.7 times as productive as conventional wells, with initial production of approximately 80,000 mcf per month compared to 30,000 mcf per month.

This simulation shows how, on a per-well basis, unconventional drilling is 2.7 times as responsive as conventional wells. However, the typical number of wells drilled each month has also changed during the shale revolution. The next simulation further accounts for these changes over time.

### 4.4.6 Simulation 2: Simulated Price Responses Over Time

The second set of simulations includes a separate simulation for each year from 2000 through 2015. For each simulation, we use year-specific values for well productivity, oil and gas prices, and baseline drilling activity. This reflects how all of those mar-
ket factors have dramatically changed during the shale revolution: higher per-well productivity, lower gas prices, and fewer wells drilled (despite higher production).

For each simulated year, we use that year’s average oil and gas prices. We simulate a $1 per MMBTU shock to gas prices. We scale the production profiles shown in Figure 4.6 to match each year’s average initial production, by well type. We specify baseline drilling activity using the average nationwide monthly gas spuds. We use nationwide spuds (as opposed to Texas only) to approximately extrapolate to nationwide unconventional/conventional drilling shares.

The results are shown in Figure 4.8. Differing somewhat from the previous simulation, these figures show the change in wells drilled (left panel) and gas production (right panel) from both well types combined. This figure only shows the response during the first five years of the simulation to focus on the short- and medium-run responses. The lines are shaded in gray scale in proportion to the share of baseline drilling activity that was dominated by conventional (lighter) or unconventional (darker) drilling in the simulation year.

The left panel shows a rise in the response of drilling activity to a $1 price shock between the pre-shale era (e.g., 2000-2005) to recent years. The drilling response has moved up and down over the years: for example, the drilling responses in 2013-2015 are smaller than the responses in 2010-2012. This is due to the falling baseline number of wells drilled during this period, as illustrated in Figure 4.4. Despite falling drilling activity, U.S. natural gas production has risen substantially during this period, in large part due to rising productivity. This illustrates an offsetting impact of rising productivity: the industry can produce more gas with fewer wells. For example, the response in the number of wells drilled is approximately the same in 2012 and 2003, but the response in gas produced is nearly twice as large in 2012.

This is further illustrated in the right panel of Figure 4.8, which shows a clear trend of a rising production responsiveness over time, despite the weaker trend in

---

61 We shift from a percentage shock in the previous simulation to a dollar-valued shock to reflect the sharply changing price environment.

the response of wells drilled. This contrast is due to rising productivity, causing a given change in wells drilled to result in increasing gas production.

The change in the overall responsiveness of gas supply due to the shale revolution is large. The average five-year response for the 2010-2015 “shale era” simulations is about 5.4 bcf/d, almost all of which comes from unconventional wells. This is about three times larger than the average response for the 2000-2005 “pre-shale” simulations of about 1.9 bcf/d, which features relatively little unconventional activity.

![Figure 4.8: Change in Gas Wells Beginning Production and Natural Gas Produced from Those Wells, Following a $1 per MMBTU Price Shock](image)

**Sources:** Authors’ calculations based on data from Drillinginfo, Bloomberg, and EIA

### 4.5 Conclusion

We empirically analyze drilling and production from approximately 62,000 gas wells in Texas from 2000 to 2015 to examine whether unconventional gas supply is in fact more responsive to price changes than conventional sources, as has been widely conjectured. We consider separate stages of the natural gas extraction process: drilling investment, time to initial production, and the profile of output over time. We find that neither production from existing wells nor completion times respond strongly to price changes. Rather, the important margin for supply response is drilling investment. We estimate a gas drilling response with respect to gas prices of approximately 63 For all simulations after 2010, more than 90 percent of the incremental gas supply comes from unconventional wells.
0.9, finding no evidence that this response is different for unconventional versus conventional gas wells.

We also find important differences between conventional and unconventional wells. While unconventional wells tend to take longer to reach production, they produce much more gas per well than conventional wells and have much lower percent variation in production, consistent with the notion of a manufacturing process. The faster flow rate per well turns out to be the primary margin by which aggregate supply from unconventional gas production is more price-responsive than conventional reservoirs. In particular, on a per-well basis, we find an approximately 3-fold greater responsiveness of unconventional gas supply to price changes compared to conventional gas, due to greater well productivity.

Additionally accounting for changes to the level and composition of drilling activity, the responsiveness of natural gas supply is about three times as price responsive during the “shale era” of 2010-2015 compared to the “pre-shale era” of 2000-2005. We illustrate that the distinctions between the stages of production (drilling, completion, and production) are key to understanding price responsiveness. These distinctions are critical in an industry where drilling rigs are a major cost factor and the total number of operating rigs is slow to change. Among other important results, this research demonstrates why simply counting wells drilled or rigs operating is no longer sufficient to gauge changes in gas supply, without also measuring heterogeneity in well productivity.
Chapter 5

Conclusion

Each of the preceding chapters has implications for environmental and energy policy.

In Chapter 2, I showed that behavioral factors drive consumer responses: consumer awareness and information are key to inducing responses, and marginal prices have minimal effects. This has implications for the design of effective electricity pricing; namely, policymakers should focus on ensuring that consumers understand their pricing mechanism, rather than attempting to simply target policies or to make marginal adjustment to the pricing regimes themselves. This can help amplify consumers’ responses to TOU pricing policies, more effectively reducing peak load, reducing costs and (in some cases) improving environmental outcomes as well.

Chapter 3 illustrated an under-appreciated value of quantity-based regulation over price-based regulation: firms’ anticipations of policy improvements are quickly factored into allowance markets, allowing for a more rapid incorporation of new information about a policy’s costs and benefits. On the other hand, this factor can be undesirable when policies are expected to deteriorate over time. This result has implications for designing environmental regulatory policies, and it can help us better interpret the impacts of the significant price volatility exhibited in allowance markets, such as the SO$_2$ program.

Chapter 4 estimates time-varying supply elasticities for natural gas, which is
important to understand as we move toward relying more heavily on natural gas to fuel electricity generation. Such a shift appears inevitable, prodded on by market forces and expected climate policies. We find a much larger price response of gas supply, which would diminish the impacts of demand shocks on gas price levels and volatility. Altogether, this suggests that a continuing shift away from coal-fired generation is a less risky proposition than it would be absent the shale boom.
Appendix A

Appendix to Chapter 2

A.1 Experimental Details

Figures A.1, A.2, A.3, and A.4 depict the invitation letter, energy usage statement, the refrigerator magnet/stickers, and the in-home electricity monitor.

A.2 Robustness Checks

A.2.1 Pre-trend Check

To help assess the validity of the parallel trends assumption, Figure A.5 shows average monthly peak-period consumption for both treatment and control groups. The time series pattern of consumption appears similar for the treatment and control groups, suggesting no cause for concern of differential time trends in energy consumption between treatment and control groups. The pre-trends for the groups also look similar for nighttime consumption and daytime off-peak consumption (not shown).
Mr AB Sample
123 Sample Road
Sample Town
Sample County

17th November 2009

The National Smart Meter Plan. Make the smart move.

Dear Mr Sample,

I would like to invite you to join the User Trial of the National Smart Meter Plan due to commence shortly. If you choose to take part you will receive £50 credited to your ESB bill as a "thank you" at the end of the Trial.

What is a Smart Meter?
A Smart Meter is a new electricity meter that:
- Helps you make savings on your bill by helping you reduce the amount of electricity you use
- Provides you with more information on what electricity you use
- Reduces estimated bills
- Assists you in being energy efficient

In this trial you will be given a Smart Meter. The purpose of the Trial is to see how Smart Meters can help customers reduce the amount of electricity they use. Smart Meters are already in use in other countries and following the trial, the feedback will be evaluated before any award.

What happens if you choose to participate?
- ESB Networks will replace your existing electricity meter with a Smart Meter.
- ESB Customer Supply will provide you with information at the start of the trial and at regular intervals throughout.
- You will receive £50 credited to your ESB bill at the end of the trial as a "thank you" for completing the trial.

During the User Trial a number of different ways of improving your energy efficiency will be tested, including:
- Time of Use Tariffs (different prices for different times during the day and night).
- Feedback and other information will offer you the opportunity to better manage and reduce your electricity bill.

How do you sign up to take part in the Smart Meter User Trial?
- Complete and return the leaflet at the bottom of this letter.
- Or phone us on the Lo Call number 1850 21 18 50.
- Or email us at smartmeter@esb.ie quoting your ESB Customer Supply account number and your name and address.

As this User Trial is limited to 5,000 places nationally, we would ask you to let us know as soon as possible if you wish to participate.

Please note, throughout the trial, all information will remain confidential to you and to ESB Customer Supply.

Want to find out more?
Please read through the enclosed leaflet or, if you prefer, call us on 1850 21 18 50 or email us at smartmeter@esb.ie

Yours sincerely,

Pat Fagan
General Manager
ESB Customer Supply

The National Smart Meter Plan is managed by the Commission for Energy Regulation with the support of the Department of Communications, Energy & Natural Resources, Data Protection, Standards, and the electricity industry in Ireland.

Customer Supply

Customer Name: Mr AB Sample
Account No: 00000000

Address: 123 Sample Road
Sample Town
Sample County

I agree to participate in the User Trial of The National Smart Meter Plan and to receive information in relation to this Plan from ESB Customer Supply. (Please tick)

My contact number is

My email address is

We need your contact number so that ESB Networks can call you to arrange an appointment to install your Smart Meter. Your contact number will be used for account management purposes only. Please note, throughout the trial, all information will remain confidential to you and to ESB Customer Supply.

Figure A.1: Example Invitation Letter
A.2.2 Placebo Test: Holidays and Weekends

Figure A.6 shows the results of a placebo test, in which I re-estimate equation (2.1) for periods when time-of-use pricing was not active: holidays and weekends. It shows little average treatment effect. The is a small but significant reduction during peak hours (when prices are normally high on a weekday) and a small but significant increase between 12:00am and 12:30am (when prices are normally be low on week-
days). This is likely explained by some combination of consumer confusion about holiday/weekend pricing, habit formation, and scheduled automation of home appliances.\footnote{Fowlie et al. (2017) found an analogous effect in an experiment from California.} However, the fact that the treatment effect mostly disappears relative to Figure 2.2 confirms that the key results are indeed real.

These results also suggest that the information treatments did not on their own serve to significantly reduce consumption. If they were effective, we should see treatment effects when time-of-use pricing is not applied, regardless of the time of day.
No such pattern appears in Figure A.6, with a few small but marginally significant exceptions. This suggests that the treatment effect is driven by reaction to time-of-use pricing, not simply information provision. Of course, this does not mean that information provision is irrelevant. As the results in section 2.5 show, information provision can amplify the effects of time-of-use pricing, even if it is not particularly effective on its own.

A.2.3 Time-Varying in Responses to TOU Pricing

In this section, I assess whether there are time-varying differences in how households respond to TOU pricing. To estimate this, I estimate the following variant of equation
Notes: Error bars represent 95% confidence intervals. Solid dots are statistically significant at the 95% level; empty dots are not. The horizontal placement of points indicates the middle of time period (e.g., the first point, located at the x-coordinate of approximately 0.25, is the half hour spanning 12:00am – 12:30am, and the last point laying between the two red vertical lines spanning 18:30 – 19:00 (6:30pm – 7:00pm), which is the last half-hour of the peak period). Peak periods are shown for comparison with Figure 2.2, but in fact rates during holidays and weekends were actually fixed at the day rate value. Only days without active time-of-use pricing (weekends and holidays) are included. For each point estimate, N = 492,850, except for 12:30am – 1:00am and 1:00am – 1:30am where some hours are missing due to daylight savings time, for which N = 489,844. Standard errors are two-way clustered at the household and week-of-sample levels.

Figure A.6: Placebo Test — Estimated Treatment Effects on Holidays and Weekends

(2.1):

$$\ln(Y_{i,t}) = \sum_{w=1}^{53} \eta_w W_{i,w,t} + \alpha_i + \lambda_w + \epsilon_{i,t}. \quad (A.1)$$

where only peak periods are included in the regression (5:00pm to 7:00pm). This specification allows for different treatment effects for each week of the year. Figure A.7 shows the estimated $\eta_w$ values along with 95 percent confidence intervals. The estimated treatment effects do not show strong seasonality over time. The only
notable differences are that the effect appears to be a slightly bigger in the initial weeks of the trial, and perhaps slightly smaller during the final weeks. The latter effect could be due to two factors: either changes in responsiveness near the end of the trial, and/or a lower likelihood of responding during the December holiday season when outdoor demand for lighting increases.

One caveat must be made about the weekly estimates. The baseline period begins July of 2009 (during week 28). Therefore, the estimated weekly effects during weeks 1 to 27 (January-June) primarily rely on comparisons against the July-December baseline period. Hence, the effects estimated for the first half of the year (weeks 1 to 28) depend on somewhat different assumptions than the estimates for the second half of the year (weeks 29 to 53). In each case, the identifying assumption is the parallel trends assumption. The caveat is that the parallel trends assumption is more plausible in the second half of the year. For example the estimates for second half of the year compare, e.g., week 52 of 2010 to week 52 of 2009. The parallel trends assumption is more plausible in that case than it is for the first half of the year, when the identification comes from, e.g., comparing week 1 of 2010 to the average of weeks 28-53 of 2009. The identifying assumption of parallel trends would be satisfied if the mean change from weeks 28-53 of 2010 to week 1 of 2009 would be the same across the control and treatment groups, had it been untreated. While this is a stronger assumption, it is nonetheless plausible, lending credence to the estimates in Figure A.7.

A.2.4 Causal Tree with Propensity Score Weighting

The difference-in-differences estimator is ideally removes time-invariant differences between the households, which is important in the presence of imbalance between the treatment and control groups. As a sensitivity, I present alternative regression tree results using another common method to adjust for imbalance: propensity score
weighting. Specifically, I re-estimate the causal tree algorithm using inverse propensity score weights, where the weights are the inverse of the fitted values of a logistic regression of treatment status as a function of observables (i.e., the logistic analogue of the specification shown in Table 2.3). The resulting tree is shown in Figure A.8. It strongly resembles the results from the primary specification: awareness is the key variable, followed by baseline consumption levels and information treatment. A few of the splits that appear in the primary specification do not appear in this tree, but they are not key to the primary findings.
Aware of Tariff Change?
Baseline Average Peak Consumption $\geq 0.12$ kWh (5th percentile)
Info. Treatment: In-Home Display?

yes
no

TE: $-8.9\%$
SE: (0.9\%)
n: 2328

TE: $-10.4\%$
SE: (1.4\%)
n: 2029

TE: $-11.4\%$
SE: (1.3\%)
n: 1933

TE: $-14.7\%$
SE: (1.6\%)
n: 474

TE: $-10.3\%$
SE: (1.3\%)
n: 1459

TE: $8.6\%$
SE: (12.4\%)
n: 96

TE: $-2.3\%$
SE: (1.7\%)
n: 299

Notes: TE = Treatment effect; SE = standard error; n = number of treated observations in node. If the condition is satisfied, proceed down left branch. This figure was generated using the \texttt{rpart.plot} package (Milborrow 2016) in R.

**Figure A.8:** Propensity Tree Treatment Effects During Peak Periods

### A.2.5 Honest Tree

The causal tree results in the body of the paper (shown in Figure 2.4 were estimated using the full dataset. While the size of the tree was determined through cross-validation, the same dataset was used to estimate the nodes of the tree (i.e., split the data into subgroups) and estimate the treatment effects in those nodes. Athey and Imbens (2016) caution that building one’s model and estimating parameters on the same data can possibly overstate the magnitude of the estimates, even if one regularizes with cross-validation. This is true of any flexible non-parametric estimation process, and it is distinct from standard overfitting (which is addressed through regularization and cross-validation).
To solve this problem, Athey and Imbens (2016) suggest a further extension in which one subsample of the data is used to build the tree and another one is used to estimate the treatment effects in each subgroup. In particular, the researcher first divides the dataset in half. Then, she uses one sample of the data (say, half of the data) to estimate the structure of the tree (the nodes, or division into subgroups). The researcher uses cross-validation within this sample to determine the optimal size of the tree. So far, this process is identical to what was described in the body of the paper, but using only a subsample of the full dataset. The difference is that the researcher then ignores the estimated treatment effects in this half of the data. Instead, the researcher uses the structure of the tree determined on the first half of the data, but estimates the treatment effects in the corresponding nodes of the tree using only the yet-unused second half of the data. This provides a further safeguard that any resulting heterogeneity in estimated treatment effects is real, not an artifact of the data that survives the cross-validation process. Athey and Imbens (2016) call this an “honest” causal tree, finding it has superior confidence interval coverage in simulations.

As a robustness check, I estimate an honest causal tree. I divide the data in half and estimate the structure of the tree using cross-validation in that half. Holding the tree’s structure fixed, I then use the other half of the dataset to estimate the treatment effects in each subgroup determined in the first step. The results for peak periods are shown in Figure A.9. The results from the honest tree confirm the main results: aware households are exhibit much larger reductions in peak consumption (-10% versus 1.4%).\(^2\) After this, baseline consumption and information treatment matter, with larger households and households with the in-home display responding more.

\(^2\) I also estimated honest trees for nighttime and daytime off-peak periods. Consistent with the results in the body, I find little-to-no heterogeneity in treatment effects during those periods.
This tree has two differences compared to the estimates on full dataset. First, the splits by baseline consumption are somewhat different: the honest tree chooses a somewhat higher cutoff for dividing aware households (0.22 versus 0.12 kWh), and it does not split unaware households at all. This difference is minor. Second, of the information treatments, only the in-home display split survives cross validation on this smaller sample. Nonetheless, the branches of the honest tree pruned by cross validation are the same as in the full-data tree. These branches (not shown) include a split on baseline consumption for unaware households and two splits below the in-home display node for the monthly bill treatment followed by the overall load reduction incentive treatment.

**Figure A.9**: Honest Tree Treatment Effects During Peak Periods
A.3 Confirming No Heterogeneity on Appliances and Temperature

The causal trees find no robust sources of heterogeneity in average treatment effects by household appliance usage (or other household demographics). In this section, I confirm this by estimating effects separately by appliance ownership. I also estimate differences by temperature bins to consider the possibility that the lack of any clear differences by appliance ownership is masking the existence differences that occur only at extreme temperatures. These effects are estimated using the same panel regression method, separately by appliance ownership and temperature bin. Temperature is measured as the wet bulb temperature in Phoenix Park, Dublin. I do not have location data at the household level, but the households in the sample are disproportionately located in Dublin (according to Commission for Energy Regulation (2011a), more than 25% of households in the experiment are in Dublin). Temperature is binned into eight 3°C-wide intervals ranging from -4°C to 20°C, a histogram of which is shown in Figure A.10.

The temperature-binned treatment effects are shown in Figures A.11 and A.12. Figure A.11 shows no clear difference in the average effect across temperature bins. Figure A.12 shows this separately for homes with and without a variety of common appliances: electric home heating, electric water heating, electric cook stoves, dishwashers, and tumble dryers. There are generally no major differences in responses across these groups. The one exception is in the top right panel of Figure A.12, which shows a marginally significant difference of water heating, but only at very low temperatures. Nonetheless, this is driven by a smaller response (relative to warmer temperatures) among households without electric water heat, and not by a larger response (relative to warmer temperatures) among households with electric water heat. Indeed, among households with electric water heat, the response is no
larger at cold temperatures than at warm ones. This is the opposite of what one might expect if the explanation was that those with electric water heaters respond more at cold temperatures. Altogether, these results suggest that the average response cannot be attributed to solely one margin. Instead, households appear to be responding in many ways. The evidence points towards water heating as perhaps a somewhat more important margin than others, but the statistical significance of this effect is marginal.

![Figure A.10: Histogram of Temperature Bins](image-url)
A.4 Overview of Regression Trees and Cross Validation

This section provides an overview of regression trees and cross-validation for readers with little background in these methods.

Regression trees, also referred to as classification and regression trees (CART), recursively partition the covariate space into subsets called “branches.” For example, a branch could divide the data into households with a college education versus those without. A branch could also divide the data along continuous variables, such as splitting the data according to whether energy consumption is above or below some threshold (say, 0.2 kWh), where the threshold is chosen to maximize model fit. Within a branch, the predicted outcome variable is simply the sample average of the outcome variable ($Y_i$) for observations in that branch. At each step, the variables chosen to split are those that best reduce the in-sample model fit. The algorithm continues to divide the subsetted data along new branches until additional splitting
would not sufficiently increase the model fit. The end result is that the data is divided into non-overlapping hyper-rectangles that divide observables into discrete subsets called “leaves”. In standard regression trees, every observation then falls into exactly one leaf with its own predicted value, equal to the simple average of the outcome variable for all observations in that leaf.

Trees that are grown in an unrestricted manner will generally overfit the data,
leading to poor predictions out of sample. To mitigate this, trees are “pruned” using cross-validation. This involves removing branches to produce a smaller, parsimonious tree. The optimal tree size is determined by cross validation.

In cross validation, the data is randomly divided into $K$ sets (typically 5-10) of approximately equal size. For each set, the modeler sets aside that data and uses the remaining observations (called the training set) to fit trees of varying sizes (from a null model of no splits at all to the fully grown tree). Each model is used to predict the outcome variable in the hold-out set (also called the test set). The out-of-sample predictive performance of each model is assessed by computing the root mean square error (RMSE) of the predicted outcome value versus its true value in the test set. This process is repeated for each of the $K$ hold-out sets, and the $K$ test RMSEs for each model are then averaged, providing a measure of out-of-sample performance of models of varying complexity. Highly complex models tend to overfit the data, leading to large mean test RMSEs and are generally rejected by cross-validation. The model with the optimal level of complexity according to test RMSEs is then chosen. Finally, this model is fit on the full data. This leads to a parsimonious tree that is simple to interpret and has superior out-of-sample prediction properties.

**A.5 Consumer Surplus**

In this section, I consider the costs to consumers and capacity-related benefits of switching from flat-rate pricing to time-of-use. The capacity benefit is the lesser need for utilities to build expensive peaking capacity, and it represents the primary benefit of TOU pricing.³ The primary costs of the policy are the forgone consumption

³ Another benefit is the reduced generation cost during peak periods, of course offset by increased generation costs resulting from increased consumption off-peak. I do not estimate energy savings because they are generally much smaller than the capacity benefit. For example, Fowlie et al. (2017) estimates that the energy-related benefits of TOU pricing in California are less than one-fifth the size of the capacity benefits.
by households who reduce their peak consumption in the face of higher prices (as well as any adjustment costs), although this is partially offset by households enjoying somewhat more consumption at lower prices during off-peak hours.

The lost consumer surplus (CS) from TOU pricing is estimated to be fairly small when estimated as the “triangle under the demand curve” method: 3€ on average per household per year on average, which ranges from 2€ without an IHD to 4€ with an IHD (approximately $2 to $5 per household per year, converted at the 2010 average exchange rate of $1.33 per € from Bloomberg L.P.). These losses are relatively small, but they also ignore any program implementation costs, including the purchase cost of an IHD.4

These estimated losses of 3€ to 4€ represent the “triangle cost”, which is the sum of the hour-by-hour products of price changes and consumption changes, divided by two. The triangle cost excludes the transfer from households to suppliers, in which households pay higher prices on inframarginal units of electricity consumption. Excluding such a transfer is typical because it is does not represent a true social cost. However, to gain an more nuanced view of the effect on consumers, I also estimate the full change in consumer surplus, which includes both the “triangle cost” as well as the transfer (the change in price multiplied by the \textit{ex post} consumption). The effects on consumer surplus are estimated to be 18€ to 19€ per household per year. These effects are larger than the triangle cost, but they are nonetheless represent only a couple percentage points of the approximately 900€ average annual Irish expenditure on electricity.5.

Capacity benefits are estimated to be a one-time benefit of 100€ per household per

4 Unfortunately I do not have access to the cost of the IHDs used in this experiment. For reference, more advanced IHDs sold in the United States typically cost about $250 (see e.g., Bollinger and Hartmann (2016) and www.nest.com, which as of March 29, 2017 priced Nest thermostats at $249).

in Ireland on average, ranging from €50 without an IHD to €130 with an IHD. The estimate of €50 assumes a weak information treatment (the bi-monthly bill), whereas the estimate of €130 assumes a strong one (an in-home electricity display, or IHD). The difference of €80 can be interpreted as the average capacity benefit of the IHD.  

More generally, capacity benefits are likely to be larger in regions where households consume more electricity such as the United States. This is because a fixed percentage reduction in peak demand will reduce more capacity (in kW) in high-consumption regions. Applying the same percentage reductions found in this experiment to peak demand levels consistent with the United States results in estimated benefits on the order of $170 to $435 per household (without and with an IHD, respectively, converted at the average 2010 exchange rate of $1.33). In this case, the value of the IHD (the $265 difference) is estimated to be larger because of the larger number of kilowatts reduced (even though the percentage difference is constant).

A.5.1 Costs: Effect on Consumers

Table A.1 shows estimates of the effect of TOU pricing on consumers. Three metrics are reported: the change in surplus excluding this transfer (reflecting true social cost), the change in consumer bills, and the change in consumer surplus including the transfer to suppliers. The first value is most appropriate for a cost-benefit analysis, but the others are informative for consumer impacts. These different effects

---

6 Annualizing these one-time benefits at a discount rate of 8% (derived below) suggests annualized benefits of €4 to €10 per household per year, which are larger than the estimated costs of €2 and €4, respectively.

7 All estimates I present ignore the lump-sum payments given to households for their participation in the experiment. I ignore these payments because I want to isolate the impact of TOU pricing itself. These payments ranged from €55 to €140 in total: €50 compensation for participating in the surveys plus a payment of either €30, €50, €70, or €90 to households in the A, B, C, or D tariff groups. Accounting for these payments, the vast majority of households benefited from their participation in the experiment.
are illustrated conceptually by Figure A.13. The figure illustrates a price increase $(p_0 \rightarrow p_1)$, but the computations are analogous for a price decrease $(p_1 \rightarrow p_0)$, albeit with impacts of the opposite sign.

All effects were simulated using the heterogeneous impacts on peak pricing presented in Figure 2.4. For each half-hour in the treatment period, I simulate each treated household’s counterfactual consumption (i.e., how much they would have consumed if they were not treated, using the relevant treatment effect given the household’s characteristics, treatment group, and time of day). I then compute the change in the consumer’s half-hourly electricity costs and consumer surplus (both with and without the transfer). The policy generally raises costs during peak hours and reduces them during off-peak hours.\(^8\) These changes are summed to the annual level for each household. I then present the average of these values across households. The three columns in Table A.1 show the average across all households (first column) as well as averages for households in the weakest and strongest information treatments (bi-monthly bill and IHD, the final two columns).

The averages suggest relatively small net costs for consumers from this policy. The fact that there are net costs reflects the fact that added costs during peak hours exceed the benefits from off-peak hours on net. However, there is still substantial heterogeneity in households, with many households experiencing reduced bills due to the policy. The distributions of the annual impacts on bills are shown in Figure A.14, separately for the two chosen information treatments. This reveals substantial heterogeneity, showing that TOU pricing could reduce bills for some households while increasing them for others (as much as 50%). Whether bills increase or decrease depends on the time profile of their consumption and their price responsiveness.

For a household to reduce its bill, it must reduce its electricity use during peak

\(^8\) For this reason, it is possible for the net cost to be negative, with the lower prices off-peak more than offsetting the higher on-peak prices.
hours, forgoing consumption that has value to households. This explains why IHD households exhibit larger losses of consumer surplus (including the transfer) despite having lower bills: they are forgoing consumption, which is costly but shows up as a savings on their bills. It also explains the lower loss in consumer surplus (excluding transfer) for the bi-monthly bill group: they are not reducing as much, implying smaller lost value of foregone consumption. Their consumer surplus losses are mostly transfers to the utility.
Table A.1: Effect on Treated Consumers during Treatment Period (Jan-Dec 2010)

<table>
<thead>
<tr>
<th>Loss in CS Excluding Transfer (b)</th>
<th>Average Household</th>
<th>Bi-Monthly Bill</th>
<th>IHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss in CS Including Transfer (a + b)</td>
<td>€18.64</td>
<td>€19.18</td>
<td>€17.66</td>
</tr>
<tr>
<td>Change in Bill (a – d)</td>
<td>€5.22</td>
<td>€8.61</td>
<td>€1.38</td>
</tr>
</tbody>
</table>

Notes: Letters in row labels refer to labeled areas in Figure A.13. For example, loss in CS including the transfer (a + b) is much larger than the increase in consumer bills (a – d) primarily because it also includes the value of the forgone electricity consumption (d).

A.5.2 Benefits: Capacity Savings from Reduced Peak Load

In this section, I estimate the avoided construction costs (“capacity benefit”) associated with the effects from the treatments tested in this experiment. It should be noted that the following estimates are only approximate and should be considered to be “back of the envelope” due to the uncertainties in the various inputs and the rather strong assumptions required.
For purposes of external application, I estimate the per-household capacity benefit. I construct estimates based on three alternative estimated peak-period treatment effects: the average treatment effect (-8.9%), the treatment effect for the bi-monthly bill group (-4.8%), and the treatment effect for the IHD group (-12.5%). This distinction allows for an estimation of the value added by the IHD. I also construct each estimate for Ireland as well as for the United States, where households have higher energy consumption and hence more potential for peak reductions. However, one should be cautious in interpreting the estimated U.S. benefits because it involves extrapolating my results from Ireland to a very different country, where electricity

9 These treatment effects differ from those shown in Figure 2.4 because those also condition on awareness and baseline energy consumption.
demand is driven by different factors. Nevertheless, some literature suggests comparable effects of dynamic pricing in the United States in a variety of settings, not only residential but also commercial and industrial (e.g., Jessoe and Rapson 2014; Bollinger and Hartmann 2016; Blonz 2016), so these estimates are likely of the correct order of magnitude.

The results are shown in Table A.2. All values have been rounded to the nearest €5 (or nearest $5 in the case of the United States). They suggest substantial benefits on the order of €50 to €130 in Ireland, depending whether an IHD is included. The corresponding figures for the United States are $170 to $435 per household, and on the region (Ireland versus United States). The IHD more than doubles the benefits because the estimated reduction on peak load is much larger with an IHD. The effects are larger in the United States—even on a per-household basis—because American households use more than twice as much electricity as Irish ones do, meaning equivalent percentage changes have larger effects in watts.

The additional capacity savings an IHD generates (€80 in Ireland and $265 in the United States) do not appear to justify the IHD’s cost in Ireland but may narrowly do so in the United States. While I do not have access to the cost of the particular IHDs used in this experiment, similar smart thermostats and electricity monitors cost about $250 in the United States.\footnote{See e.g., Bollinger and Hartmann (2016) and www.nest.com, which as of March 29, 2017 priced Nest thermostats at $249.} However, those devices generally have more functionality than the more simple IHDs used in this experiment, likely leading to larger benefits. This suggests the added value I estimate is a lower bound for the value of more advanced devices like smart thermostats.

**Estimating Avoided Costs**

The benefits in Table A.2 are computed as the construction costs avoided due to
lower peak demand. TOU pricing reduces peak load, mitigating the need to build expensive peaking plants that are only activated a few times each year and otherwise sit idle. For each watt of peak load that is reduced, utilities can reduce investment in an additional watt of peaking capacity. This implies that the costs avoided due to a successful TOU pricing program are simply the product of the reduction in peak load (in watts) due to the program and the per-watt cost of building and maintaining a peaker plant.

For my estimate of the cost of capacity, I use the average of estimates from reports by the EIA and Brattle Group, resulting in an estimated cost of $1,057,000 per MW, or $1.057 per watt.\textsuperscript{11} In addition, there are fixed operation and maintenance (O&M) costs required to keep the plant in operating condition, which I estimate to be $8.76 per kW-year,\textsuperscript{12} which is approximately $86 per kW (or about $0.086 per watt) in present value assuming a discount rate of 8% and 20 year plant life,\textsuperscript{13} Adding this to the construction costs yields a total NPV avoided cost of $1.14 per watt of peak load

11 Capital costs are the average of estimated costs by EIA (Table 1, Conventional CT) and Brattle Group (Table 1, “Installed” cost). EIA costs are in 2012$ whereas Brattle estimates are in 2018$, assuming a 3% inflation rate. Therefore, I escalated EIA figures at 3% to 2018 to be comparable with Brattle’s estimates.

12 Based on EIA (ibid.), Table 1, Conventional CT, Fixed O&M, adjusted to 2018$ at 3% annually, as described in the previous footnote.

13 I use the 8% discount rate suggested by the Brattle Group (same source as above, Table 25). I assume a 20 year economic life, also from Brattle (page iv).
reduced.\textsuperscript{14} In order to compare the benefits to the costs to Irish households (which are computed in Euro), I convert this value to Euros at the average 2010 exchange rate of €0.75 per dollar, yielding a cost per watt of €0.86. For the United States calculation, I use the value in dollars.

The reduction in peak consumption from the program is also straightforward to calculate. Peak load for the entire Ireland electricity grid during the 2009-10 winter occurred at 5:30-6:00pm on Thursday, January 7, 2010. During that half hour, average household load among treated households treatment group was approximately 1.2 kW. As previously shown, the average effect on peak load is approximately -8.9%. For those with the weakest information treatment—the bi-monthly bill—the effect was -4.8%, compared to -12.5% for those who received an IHD.\textsuperscript{15}

Using these three estimated treatment effects, the estimated reductions in peak load are 0.11 kW (average), 0.06 kW (without an IHD), and 0.15 kW (with an IHD).\textsuperscript{16,17} At a value of €0.86 (derived above, and equivalent to $1.14 per watt), this implies a benefit of approximately €100 per household (≈€0.86 per watt × 110 watts) on average. For households without IHD, the benefit is €50 per household (€0.86 × 60

\textsuperscript{14} This estimate is comparable to the estimate of $1.19 per watt used in Blonz (2016).

\textsuperscript{15} This effect on peak consumption appears to be even larger during the peak day of January 7, 2010 (-22% on average that day, compared to -8.9% for the full year). To be conservative, I use the annual average treatment effect. The very large estimated treatment effect during the first few weeks of the experiment may owe to the fact that the treatment had just begun and was fresh in participants’ minds. Such large treatment effects are unlikely to be sustainable in the long run, and indeed the effect stabilized soon after around its average level, as seen in Figure A.7.

\textsuperscript{16} This is computed as the observed treated peak consumption in the sample, scaled up to the untreated counterfactual level (e.g., $\frac{1.2\ kW}{1-0.089}$ for the average effect), multiplied by the treatment effect (e.g., −0.089).

\textsuperscript{17} One may be concerned that this may overstate the reduction in peak demand. In particular, large reductions during peak periods may be offset by shifting to off-peak times. If either effect is substantial enough, TOU pricing could simply change the timing of the peak, rather than its overall level. The data suggests this is not the case, as shown by Figures 2.2 and 2.3. Figure 2.2 shows no statistically significant increase in consumption except for a small increase during 11:00pm-12:30am, which generally features the lowest load and is not at risk of becoming the new peak. Figure 2.3 shows no shifting of the peak hour on average; instead it suggests a general reduction in load in the key 5pm-10pm window. This suggests that shifting of the peak to other times is not a problem.
watts), or €130 with an IHD (€0.86 × 150 watts).

Extrapolating these figures to the United States raises several issues. First, American households may respond differently than Irish households for a variety of reasons, including differences in the drivers of peak demand and differences in income. Evidence from the United States (e.g., Blonz 2016; Jessoe and Rapson 2014) suggests effects of similar magnitude, implying this approximation is reasonable. A second difference is that Irish households use much less electricity than U.S. households do (approximately 60% less). This means that for a given percentage change in consumption, the capacity savings (in watts) would be larger for American households—perhaps more than twice as large. I account for this in my estimates through an adjustment based on U.S. electricity demand.

With these issues in mind, I present a rough order of magnitude of the potential benefits of TOU pricing to American utilities. Assuming American households respond at similar percentages as the Irish households do, the capacity savings are estimated to be $325 on average, $170 per household without an IHD, and $435 per household with an IHD. These figures are simply the Irish per-household benefits (€100, €50, and €130 respectively) converted to Euro at the 2010 exchange rate ($1.33 per Euro) and scaled by the ratio of U.S. and Irish electricity consumption, which is approximately 2.5.18

---

18 In the United States in 2015, average annual residential electricity consumption is approximately 11,270 kWh per household (based on EIA and Census data). According to the Sustainable Energy Authority of Ireland, the corresponding value for Ireland is 4,470 kWh per household. The ratio of these values is 2.5.
A.6 List of Survey Questions Used

Table A.3 displays a summary of the survey variables used in the analysis. This list shows 85 survey questions. The total number of variables used in the analysis is larger than this for several reasons. First, many of the variables are factor variables, which were converted to sets of indicator variables for each level (for example, employment status is a factor variable with seven levels). Second, questions asked in both surveys appear twice, once for each survey, to capture potential changes in status. Third, I also use variables representing treatment groups, which do not appear in the survey. Fourth, the smart meter data also provides other non-survey data. I only have listed the survey questions I use (others were incomplete).\textsuperscript{19} Variables measured on Likert scales (i.e., agree/disagree or satisfied/dissatisfied on a scale of 1-5) are treated as numeric in the analysis.

<table>
<thead>
<tr>
<th>Question</th>
<th>Variable Type</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Socio-demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1  Gender (recorded from voice)</td>
<td>Binary</td>
<td>Both surveys</td>
</tr>
<tr>
<td>2  May I ask what age you were on your last birthday? 18-25; 26-35; 36-45;</td>
<td>Ordered factor</td>
<td>Both surveys</td>
</tr>
<tr>
<td>46-55; 65+.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3  Moving on to education, which of the following best describes the level</td>
<td>Factor</td>
<td>Initial survey</td>
</tr>
<tr>
<td>of education of the chief income earner? None; Primary; Secondary to</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cert; Secondary without Cert; Third.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4  How many people in your household work for pay?</td>
<td>Integer</td>
<td>Follow-up survey</td>
</tr>
<tr>
<td>5  What is the employment status of the chief income earner in your</td>
<td>Factor</td>
<td>Both surveys</td>
</tr>
<tr>
<td>household, is he/she: An employee; Self-employed (with employees); Self-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>employed (with no employees); Unemployed (actively seeking work); Un-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>employed (not actively seeking work); Retired; Carer (Looking after</td>
<td></td>
<td></td>
</tr>
<tr>
<td>relative or family).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6  What is the occupation of the chief income earner in your household?</td>
<td>Factor</td>
<td>Both surveys</td>
</tr>
<tr>
<td>[coded to Irish social class: AB; C1; C2; DE; Farmer]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7  How many people over 15 years of age live in your home?</td>
<td>Integer</td>
<td>Both surveys</td>
</tr>
<tr>
<td>8  How many people under 15 years of age live in your home?</td>
<td>Integer</td>
<td>Both surveys</td>
</tr>
<tr>
<td>9  How many total people live in your home?</td>
<td>Integer</td>
<td>Computed</td>
</tr>
<tr>
<td>10 What best describes the people you live with? Live alone; Multiple</td>
<td>Factor</td>
<td>Both surveys</td>
</tr>
<tr>
<td>adults; Both adults and children</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Have you had to go without heating during the last 12 months through</td>
<td>Binary</td>
<td>Both surveys</td>
</tr>
<tr>
<td>lack of money?</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Housing Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 I would now like to ask some questions about your home. Which best</td>
<td>Factor</td>
<td>Initial survey</td>
</tr>
<tr>
<td>describes your home? Apartment; Bungalow; Detached; Semi-detached;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terraced.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{19} Full lists of all questions asked are available from ISSDA, here and here.
13 Approximately how old is your home? 0-5 years; 10-30 years; 30-75 years; 75+ years.  
Factor Initial survey

14 Do you own or rent your home? Own outright; Own with mortgage; Rent (public); Rent (private); Other.  
Factor Both surveys

15 And now considering energy reduction in your home please indicate the approximate proportion of light bulbs which are energy saving (or CFL)? None; about a quarter; about half; about three quarters; all.  
Ordered factor Initial survey

16 Please indicate the approximate proportion of windows in your home which are double glazed? None; about a quarter; about half; about three quarters; all.  
Ordered factor Initial survey

17 Is your attic insulated and if so when was the insulation fitted?  
Ordered factor Initial survey

18 Returning to heating your home, in your opinion, is your home kept adequately warm?  
Binary Both surveys

19-28 Over the last twelve months have you done any of the following?  

19 Added double glazing to some or all of your windows  
Binary Follow-up survey

20 Installed insulation to your home (attic or walls)  
Binary Follow-up survey

21 Replaced appliances with A rated ones  
Binary Follow-up survey

22 Fitted a new lagging jacket on your hot water tank  
Binary Follow-up survey

23 Fitted other energy saving devices - such as usage monitors  
Binary Follow-up survey

24 Added solar panels  
Binary Follow-up survey

25 Added draught-proofing to your doors or windows  
Binary Follow-up survey

26 Replaced a central heating boiler with a more efficient one  
Binary Follow-up survey

27 Added thermostatic controls to radiators so that their temperature could be turned down when the room is not in use  
Binary Follow-up survey

28 None of these  
Binary Follow-up survey

Appliances & Electronics

29 Do you have internet access in your home?  
Binary Initial survey

30 How many electric showers are in your home?  
Integer Computed

31 Do you have a timer to control when your hot water/immersion heater comes on and goes off?  
Binary Initial survey

32 Does your hot water tank have a lagging jacket?  
Binary Initial survey

33 Which of the following best describes how you cook in your home? Electric cooker; Gas cooker; Oil fired cooker; Solid fuel cooker; Microwave; Don’t cook.  
Factor Both surveys

34 Electricity (electric central heating/storage heating)  
Binary Both surveys

35 Electricity (plug in heaters)  
Binary Both surveys

36 Gas  
Binary Both surveys

37 Oil  
Binary Both surveys

38 Solid fuel  
Binary Both surveys

39 Renewable (e.g. solar)  
Binary Both surveys

40 Other  
Binary Both surveys

41 Do you have a timer to control when your heating comes on and goes off?  
Binary Initial survey

42 How many appliances do you own?  
Integer Computed

43-52 Please indicate how many of the following appliances you have in your home?  

43 Washing machine  
Integer Initial survey

44 Tumble dryer  
Integer Initial survey

45 Dishwasher  
Integer Initial survey

46 Electric shower (instant)  
Integer Initial survey

47 Electric shower (electric pumped from hot tank)  
Integer Initial survey

48 Electric cookers  
Integer Initial survey

49 Electric heaters (plug-in convector)  
Integer Initial survey

50 Stand alone freezer  
Integer Initial survey

166
51 Water pump Integer Initial survey
52 Immersion heater Integer Initial survey

53 How many entertainment appliances do you own and actually use?

54-58 And how many of the following entertainment appliances do you have?
Only those that are actually used should be mentioned?

54 TV's less than 21 inches Integer Initial survey
55 TV's greater than 21 inches Integer Initial survey
56 Desk-top computers Integer Initial survey
57 Lap-top computers Integer Initial survey
58 Game consoles (such as Xbox, PlayStation, or Wii). Integer Initial survey

Attitudes & Behavior

59 Do you use the internet regularly yourself? Binary Initial survey
60 Are there other people in your household that use the internet regularly? Binary Initial survey

61-67 And now, I would like to ask you a few questions about your general attitudes towards energy, electricity use and the electricity bill. [do you agree or disagree]:

61 I/we am/are interested in changing the way I/we use electricity if it reduces the bill Agree-Disagree, (1-5) Initial survey
62 I/we am/are interested in changing the way I/we use electricity if it helps the environment Agree-Disagree, (1-5) Initial survey
63 I/we can reduce my electricity bill by changing the way the people I/we live with use electricity Agree-Disagree, (1-5) Initial survey
64 I/we have already done a lot to reduce the amount of electricity I/we use Agree-Disagree, (1-5) Initial survey
65 I/we have already made changes to the way I/we live my life in order to reduce the amount of electricity we use Agree-Disagree, (1-5) Initial survey
66 I/we would like to do more to reduce electricity usage Agree-Disagree, (1-5) Initial survey
67 I/we know what I/we need to do in order to reduce electricity usage Agree-Disagree, (1-5) Initial survey

68 Does your home have a Building Energy Rating (BER) - a recently introduced scheme for rating the energy efficiency of your home? Binary Initial survey

69-72 Which of the following do you think will be benefits [of the your participation in the trial]?

69 Learn how to reduce my energy usage Binary Initial survey
70 Learn how to reduce my electricity bill Binary Initial survey
71 Do my part to help the environment by my participation Binary Initial survey
72 Do my part to make Ireland become more up to date Binary Initial survey

73-76 Thinking of what will be the main consequences of your participation in the trial, for each of the following statements, [do you agree or disagree]:

73 My household may decide to make minor changes to the way we use electricity Agree-Disagree, (1-5) Initial survey
74 My household may decide to make major changes to the way we use electricity Agree-Disagree, (1-5) Initial survey
75 My household may decide to be more aware of the amount of electricity used by appliances we own or buy Agree-Disagree, (1-5) Initial survey
76 In future, when replacing an appliance, my household may decide to choose one with a better energy rating Agree-Disagree, (1-5) Initial survey

77 How do you think that your electricity bills will change as part of the trial? Increase; No Change; Decrease.

78-83 Thinking about electricity and its use, generation and sale in the Irish context, please indicate your level of satisfaction with each of the following were 1 is very satisfied and 5 is very dissatisfied:

78 The number of suppliers competing in the market Satisfied-Dissatisfied (1-5) Initial survey
79 The percentage of electricity being generated from renewable sources Satisfied-Dissatisfied (1-5) Initial survey
<table>
<thead>
<tr>
<th>Question</th>
<th>Satisfied-Dissatisfied</th>
<th>Survey Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>The overall cost of electricity</td>
<td>Initial survey</td>
<td></td>
</tr>
<tr>
<td>The number of estimated bills received by customers</td>
<td>Initial survey</td>
<td></td>
</tr>
<tr>
<td>The opportunity to sell back extra electricity you may generate (from solar panels etc) to your electricity supplier</td>
<td>Initial survey</td>
<td></td>
</tr>
<tr>
<td>The environmental damage associated with the amount of electricity used</td>
<td>Initial survey</td>
<td></td>
</tr>
<tr>
<td>Our society needs to reduce the amount of energy we use</td>
<td>Follow-up survey</td>
<td></td>
</tr>
<tr>
<td>As part of the trial, the way you were charged for the electricity you used was changed from a single rate for all electricity to one that varies by time of day. Were you aware of this?</td>
<td>Follow-up survey</td>
<td></td>
</tr>
</tbody>
</table>
### A.7 Summary Statistics of Selected Variables

<table>
<thead>
<tr>
<th>Table A.4: Selected Summary Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Age Group: 18-25 (Indicator)</td>
</tr>
<tr>
<td>Age Group: 26-35 (Indicator)</td>
</tr>
<tr>
<td>Age Group: 36-45 (Indicator)</td>
</tr>
<tr>
<td>Age Group: 46-55 (Indicator)</td>
</tr>
<tr>
<td>Age Group: 56-65 (Indicator)</td>
</tr>
<tr>
<td>Age Group: 65+ (Indicator)</td>
</tr>
<tr>
<td>Social Class: AB, Manager/Professional (Indicator)</td>
</tr>
<tr>
<td>Social Class: C1, White collar (Indicator)</td>
</tr>
<tr>
<td>Social Class: C2, Skilled manual (Indicator)</td>
</tr>
<tr>
<td>Social Class: DE, Unskilled manual/other (Indicator)</td>
</tr>
<tr>
<td>Social Class: Farmer (Indicator)</td>
</tr>
<tr>
<td>Education: None (Indicator)</td>
</tr>
<tr>
<td>Education: Primary (Indicator)</td>
</tr>
<tr>
<td>Education: Secondary without Cert. (Indicator)</td>
</tr>
<tr>
<td>Education: Secondary to Cert. (Indicator)</td>
</tr>
<tr>
<td>Education: Third (Indicator)</td>
</tr>
<tr>
<td>Education: Refused (Indicator)</td>
</tr>
<tr>
<td>Number of Adults in Home</td>
</tr>
<tr>
<td>Number of Children Under 15 in Home</td>
</tr>
<tr>
<td>Has Children Under 15 in Home (Indicator)</td>
</tr>
<tr>
<td>Own home outright (Indicator)</td>
</tr>
<tr>
<td>Own home with mortgage (Indicator)</td>
</tr>
<tr>
<td>Home Age: Less than 5 years (Indicator)</td>
</tr>
<tr>
<td>Home Age: 5-9 years (Indicator)</td>
</tr>
<tr>
<td>Home Age: 10-19 years (Indicator)</td>
</tr>
<tr>
<td>Home Age: 30-74 years (Indicator)</td>
</tr>
<tr>
<td>Home Age: More than 75 years (Indicator)</td>
</tr>
<tr>
<td>Home Style: Apartment (Indicator)</td>
</tr>
<tr>
<td>Home Style: Bungalow (Indicator)</td>
</tr>
<tr>
<td>Home Style: Detached (Indicator)</td>
</tr>
<tr>
<td>Home Style: Semi-Detached (Indicator)</td>
</tr>
<tr>
<td>Home Style: Terraced/Townhome (Indicator)</td>
</tr>
<tr>
<td>Number of Bedrooms</td>
</tr>
<tr>
<td>Home Heat: Electric, Central (Indicator)</td>
</tr>
<tr>
<td>Home Heat: Electric, Plug-in (Indicator)</td>
</tr>
<tr>
<td>Home Heat: Oil (Indicator)</td>
</tr>
<tr>
<td>Home Heat: Gas (Indicator)</td>
</tr>
<tr>
<td>Home Heat: Solid Fuel (Indicator)</td>
</tr>
<tr>
<td>Water Heat: Electric, Central (Indicator)</td>
</tr>
<tr>
<td>Water Heat: Electric, Immersion (Indicator)</td>
</tr>
<tr>
<td>Water Heat: Oil (Indicator)</td>
</tr>
<tr>
<td>Water Heat: Gas (Indicator)</td>
</tr>
<tr>
<td>Water Heat: Solid Fuel (Indicator)</td>
</tr>
<tr>
<td>Has Immersion Water Heater (Indicator)</td>
</tr>
<tr>
<td>Washing Machine (Indicator)</td>
</tr>
<tr>
<td>Tumble Dryer (Indicator)</td>
</tr>
<tr>
<td>Dishwasher (Indicator)</td>
</tr>
<tr>
<td>Stand alone freezer (Indicator)</td>
</tr>
<tr>
<td>Cook stove type: Electric (Indicator)</td>
</tr>
<tr>
<td>Cook stove type: Gas (Indicator)</td>
</tr>
<tr>
<td>Cook stove type: Oil (Indicator)</td>
</tr>
<tr>
<td>Cook stove type: Solid (Indicator)</td>
</tr>
<tr>
<td>Number of TV’s (Less than 21 inches)</td>
</tr>
<tr>
<td>Number of TV’s (More than 21 inches)</td>
</tr>
<tr>
<td>Number of Desktop Computers</td>
</tr>
<tr>
<td>Number of Laptop Computers</td>
</tr>
<tr>
<td>Internet Access in Home (Indicator)</td>
</tr>
</tbody>
</table>

Note: Individual-specific variables (like education or age) are for the chief income earner.
Appendix B

Appendix to Chapter 3

B.1 Comparative Advantage in Two Periods

In the body, we showed that the firm’s reaction function (3.7) is \( h(\tilde{p}_t, \theta) = \hat{q} + \frac{\tilde{p}_t - c_1}{c_2}. \)

The government’s optimal policy in the first period is to tax at \( \tilde{p}_1 = c_1 \), meaning \( h(\tilde{p}_1, \theta) = \hat{q} - \frac{\theta}{c_2}. \) We also showed that under a quantity instrument, the firm produces \( \tilde{q}^{U}/2 = \hat{q} + \frac{\eta - \theta}{c_2}. \)

Now define \( \Delta^U \), the difference between the expected net benefits under a price versus a quantity instrument. If \( \Delta^U < 0 \), quantity regulation is preferred. Since there is no more uncertainty in the second period, the net benefits in that period are the same across instruments, meaning we need only consider the difference in period.
1. We now derive our primary result.

\[
\Delta U = E \left[ (B(h(p_1^U, \theta), \eta) - C(h(p_1^U, \theta), \theta)) - (B(q^U/2, \eta) - C(q^U/2, \theta)) \right] \\
= E \left[ (b_1 - c_1 + \eta - \theta)(h(p_1^U, \theta) - q^U/2) - \frac{b_2 + c_2}{2}((h(p_1^U, \theta) - q^U - (q^U/2 - \hat{q}^U)^2) \right] \\
= E \left[ (\eta - \theta) \left( \frac{\theta}{c_2} - \frac{\eta - \theta}{b_2 + c_2} \right) - \frac{b_2 + c_2}{2} \left( \frac{-\theta}{c_2} - \left( \frac{\eta - \theta}{b_2 + c_2} \right)^2 \right) \right] \\
= -E \left[ (\eta - \theta) \left( \frac{\theta}{c_2} + \frac{\eta - \theta}{b_2 + c_2} \right) + \frac{b_2 + c_2}{2} \left( \frac{-\theta}{c_2} - \left( \frac{\eta - \theta}{b_2 + c_2} \right)^2 \right) \right] \\
= -E \left[ \left( \frac{b_2 \theta \eta - b_2 \theta^2 + \eta^2 - \eta \theta}{c_2} \right) + \left( \frac{b_2 + c_2}{2c_2} \right) \theta^2 - \left( \frac{\eta - \theta}{b_2 + c_2} \right)^2 \right] \\
= -E \left[ \frac{2b_2 \theta \eta - 2b_2 \theta^2 + 2\eta^2 - 2\eta \theta + \left( \frac{b_2 + c_2}{2c_2} \right) \theta^2 - \left( \frac{\eta - \theta}{b_2 + c_2} \right)^2}{2(b_2 + c_2)} \right] \\
= -E \left[ \frac{\eta^2 + \left( \frac{b_2}{c_2} \right)^2 \theta^2 + 2b_2 \theta \eta}{2(b_2 + c_2)} \right] \\
= \frac{-(\sigma_\eta^2 + \left( \frac{b_2}{c_2} \right)^2 \sigma_\theta^2)}{2(b_2 + c_2)}.
\]

QED.

**B.2 Multiple Periods with Noise**

In this section, we prove our assertion from section 3.5 that the main result holds in multiple periods with noise. The general auto-regressive error structure for costs and benefits is given by

\[
\eta_t = \rho_\eta \eta_{t-1} + \mu_t \\
\theta_t = \rho_\theta \theta_{t-1} + \nu_t,
\]

171
and the random walk structure for the noise term:

$$\epsilon_t = \epsilon_{t-1} + \omega_t$$

with $\eta_0 = \theta_0 = \omega_0 = 0$ and $\mu_t = \nu_t = \omega_t = 0$ for $t \geq T$.

### B.2.1 Equilibrium under a Price Policy

In each period $t - 1$, the government sets a price for period $t$ to maximize its noisy view of the expected net benefits function. The first order condition is

$$E[MB_{\text{noisy}}(h(\tilde{p}_t, \theta_t), \eta_t, \epsilon_{t-1}) - MC(h(\tilde{p}_t, \theta_t), \theta_t)|t-1] = 0.$$ 

Using the firm’s reaction function (3.7), the definitions of noisy marginal benefits (3.10) and marginal costs (3.3), and canceling $b_1 = c_1$ we have

$$\rho_\eta \eta_{t-1} - b_2 \left( \frac{\tilde{p}_t - c_1 - \rho_\theta \theta_{t-1}}{c_2} \right) + \epsilon_{t-1} = \rho_\theta \theta_{t-1} + c_2 \left( \frac{\tilde{p}_t - c_1 - \rho_\theta \theta_{t-1}}{c_2} \right).$$

Re-arranging, we find the government’s optimal price of

$$\tilde{p}_t = c_1 + \frac{c_2(\rho_\eta \eta_{t-1} + \epsilon_{t-1}) + b_2 \rho_\theta \theta_{t-1}}{b_2 + c_2} = E[p^*_t|t-1] + \frac{c_2 \epsilon_{t-1}}{b_2 + c_2}.$$ 

Using the firm reaction function derived above (3.7), $h(\tilde{p}_t, \theta_t) = \hat{q} + \frac{\tilde{p}_t - c_1 - \theta_t}{c_2}$, we can find the equilibrium outcomes under the price policy:

$$q_t = h(\tilde{p}_t, \theta_t) = \hat{q} + \frac{\epsilon_{t-1} + \rho_\eta \eta_{t-1} + (b_2/c_2) \rho_\theta \theta_{t-1}}{b_2 + c_2}$$

$$= q^* + \frac{\epsilon_{t-1} - (\mu_t + (b_2/c_2) \nu_t)}{b_2 + c_2}.$$ 

That is, under a price policy, the realized outcome in each period deviates from the societal first-best level by the amount $\frac{\epsilon_{t-1} - (\mu_t + (b_2/c_2) \nu_t)}{b_2 + c_2}$ each period.
B.2.2 Equilibrium under a Quantity Policy

The banking constraint with a dynamic trading ratio is given by

\[ B_t = \tilde{R}_t B_{t-1} + \tilde{q}_t - q_t, \]

with \( B_0 = 0 \) and \( B_t = 0 \) for \( t \geq T \).

That is, we assume the problem effectively ends at period \( T \), beyond which there is no uncertainty. After that period, the government can set its optimal policy and there is no need for banking.

Our strategy is to guess the government’s rule for setting the allocation and dynamic trading ratio each period. We then show that, with these rules, this optimal quantity level is both feasible for the firm and satisfies the firm’s first order conditions for profit maximization.

Our proposed allocation rule in every period is given by

\[ \tilde{q}_t = E[q_t^*|t-1] + (\epsilon_{t-1}/(b_2 + c_2)) + \tilde{R}_t \left( q_{t-1}^* - E[q_{t-1}^*|t-2] + \frac{\epsilon_{t-1} - \epsilon_{t-2}}{b_2 + c_2} \right), \]

and the proposed dynamic trading ratio is given by

\[ \tilde{R}_t = \frac{p_t^* + (c_2/(b_2 + c_2))\epsilon_{t-1}}{\beta(E[p_t^*|t-1] + (c_2/(b_2 + c_2))\epsilon_{t-1})}, \]

This setup ensures that the intertemporal “market clearing” condition is satisfied. That is, the sum of the quantity outcomes \( q_t \) equals the sum of the allocations \( \tilde{q}_t \), appropriately adjusted for the trading ratios. This can be seen by considering the banking rule at the terminal period, \( B_T = \tilde{R}_T B_{T-1} + \tilde{q}_T - q_T \), and repeatedly substituting in for \( B_{T-1} \) with the lagged value of the banking rule, yielding

\[ B_T = \left( \prod_{j=1}^{T} \tilde{R}_j \right) B_0 + (\tilde{q}_T + \sum_{t=1}^{T} \prod_{j=t}^{T} \tilde{R}_j \tilde{q}_{t-1}) - (q_T + \sum_{t=1}^{T} \prod_{j=t}^{T} \tilde{R}_j q_{t-1}). \]

Since \( B_T = B_0 = 0 \), this implies the clearing condition:

\[ \tilde{q}_T + \sum_{t=1}^{T} \prod_{j=t}^{T} \tilde{R}_j \tilde{q}_{t-1} = q_T + \sum_{t=1}^{T} \prod_{j=t}^{T} \tilde{R}_j q_{t-1}. \]
for $t > 1$ where $p_t^* = c_1 + \frac{b_2\theta_t + c_2\eta_t}{b_2 + c_2}$ is the first-best (i.e., excluding noise) price corresponding to $q_t^*$.\(^2\) We set $\tilde{R}_1 = 0$. We now show that the firm’s choice $q_t = q_t^* + \epsilon_t/(b_2 + c_2)$ (implying $p_t = p_t^* + (c_2/(b_2 + c_2))\epsilon_t$) is both feasible given the terminal banking condition $B_T = 0$ and satisfies the firm’s first-order condition for profit maximization (the arbitrage condition $p_t = \beta\tilde{R}_{t+1}E[p_{t+1}|t]$).

To see feasibility note that this choice of $q_t$ yields $B_1 = -(q_1^* - E[q_1^*|0] + \epsilon_1/(b_2 + c_2))$ assuming $\epsilon_0 = \theta_0 = \eta_0 = B_0 = 0$ and more generally

$$B_t = -\left(q_t^* - E[q_t^*|t-1] + \frac{\epsilon_t - \epsilon_{t-1}}{b_2 + c_2}\right)$$

for $t < T$. This can be shown by induction. That is, given this value for $B_t$, $q_t = q_t^* + \frac{\epsilon_t}{b_2 + c_2}$, and the allocation and banking rules shown above, we can show that the relationship holds for $t + 1$:

$$B_{t+1} = \tilde{R}_{t+1}B_t + \tilde{q}_{t+1} - q_{t+1}$$

$$= \tilde{R}_{t+1}B_t + \tilde{R}_{t+1}\left(q_t^* - E[q_t^*|t-1] + \frac{\epsilon_t - \epsilon_{t-1}}{b_2 + c_2}\right) + E[q_{t+1}^*|t] - q_{t+1}^* - \frac{\epsilon_{t+1} - \epsilon_t}{b_2 + c_2}$$

$$= -\left(q_{t+1}^* - E[q_{t+1}^*|t] + \frac{\epsilon_{t+1} - \epsilon_t}{b_2 + c_2}\right)$$

as desired. Assuming $\omega_T = \mu_T = \nu_T = 0$, we have $B_T = 0$. Thus, $q_t = q_t^* + \epsilon_t/(b_2 + c_2)$ is feasible for the firm.

We now show that picking $q_t = q_t^* + \epsilon_t/(b_2 + c_2)$ satisfies the first-order condition for the firm. Namely, each period the firm faces:

$$V_t(B_{t-1}, \theta_{t-1}, \eta_{t-1}, \epsilon_{t-1}, \nu_t, \mu_t, \omega_t) = \max_{B_t} -C(\tilde{R}_tB_{t-1} + \tilde{q}_t - B_t, \theta_t) + \beta E[V_{t+1}(B_t, \theta_t, \eta_t, \epsilon_t, \nu_{t+1}, \mu_{t+1}, \omega_{t+1})|t],$$

\(^2\) Written as a function of the shocks, the following gives an equivalent representation:

$$\tilde{R}_t = \frac{c_1 + \frac{b_2\theta_{t-1} + c_2\eta_{t-1} + c_2\epsilon_{t-1}}{b_2 + c_2}}{\beta(c_1 + \frac{b_2\theta_{t-1} + c_2\eta_{t-1} + c_2\epsilon_{t-1}}{b_2 + c_2})}$$
where we have replaced $q_t$ with $B_t$ as the choice variable to simplify taking derivatives. The first-order condition is then

$$MC(q_t, \theta_t) = -\beta E \left[ \frac{\partial V_{t+1}(B_t, \theta_t, \eta_t, \epsilon_t, \nu_{t+1}, \mu_{t+1}, \omega_{t+1})}{\partial B_t} \right].$$

We also have:

$$\frac{\partial V_t(B_{t-1}, \theta_{t-1}, \eta_{t-1}, \epsilon_{t-1}, \nu_t, \mu_t, \omega_t)}{\partial B_{t-1}} = -\tilde{R}_t MC(q_t, \theta_t).$$

Thus, the first order condition is that $MC(q_t, \theta_t) = \beta E[\tilde{R}_{t+1}MC(q_{t+1}, \theta_{t+1})|t]$. Given $R_{t+1}$ only depends on information at $t$, we can further simplify to

$$MC(q_t, \theta_t) = \beta \tilde{R}_{t+1} E[MC(q_{t+1}, \theta_{t+1})|t].$$

This is simply the no arbitrage condition with discounting, trading ratios, and uncertainty about future shocks. Namely, today’s marginal cost must equal the discounted, trading-ratio-adjusted, expected marginal cost next period. Given the definition of $\tilde{R}_t = (p^*_t - 1 (c_2/(b_2 + c_2)) \epsilon_{t-1}) / (\beta (E[p^*_t|t-1] + (c_2/(b_2 + c_2)) \epsilon_{t-1}))$, and using the definition of $MC(q, \theta)$ from (3.1), this condition is satisfied when $q_t = q^*_t + \epsilon_t/(b_2 + c_2)$.

Thus, $q_t = q^*_t + \epsilon_t/(b_2 + c_2)$ is feasible for the firm and satisfies the first-order conditions of both the firm.

QED.

Note that this realized outcome deviates from the societal first-best outcome by the amount $\epsilon_t/(b_2 + c_2)$ in each period.
B.2.3 Optimality of the Proposed Quantity Policy to Government

In the previous section we showed that the proposed quantity policy was feasible and optimal for the firm. Now we show that the proposed policy is optimal for the government (i.e., it leads to the government’s optimal outcome). At period $t$ the government wishes to maximize noisy net benefits based on the contemporaneous values $\theta_t$, $\eta_t$, and $\epsilon_t$. Stated differently, we find the first best outcome if the noise is taken to be part of the objective function. That is, at period $t$ the optimal outcome under noise $q_t$ must satisfy the first order condition:

$$MB_{\text{noisy}}(q_t, \eta_t, \epsilon_t) = MC(q_t, \theta_t)$$

$$b_1 + \eta_t - b_2(q_t - \hat{q}) + \epsilon_t = c_1 + \theta_t + c_2(q_t - \hat{q})$$

$$q_t = \hat{q} + \frac{\eta_t - \theta_t + \epsilon_t}{b_2 + c_2}$$

$$= q^*_t + \frac{\epsilon_t}{b_2 + c_2}$$

As shown in the previous section, this is exactly the outcome realized under the proposed quantity policy. Therefore the proposed quantity policy is optimal for the government.

QED.

B.2.4 Alternative Infinite Horizon Policy

Over an infinite horizon, we no longer have $B_T = 0$ to anchor the firm’s quantity choice among all of those satisfying

$$MC(q_t, \theta_t) = \beta \tilde{R}_{t+1} E[MC(q_{t+1}, \theta_{t+1})|t].$$
In particular, while we know \( q_t = q_t^* + \epsilon_t/(b_2 + c_2) \) satisfies this condition, we could also choose any \( q_t = q_t^* + \epsilon_t/(b_2 + c_2) - (1 - \Gamma)(p_t^*/c_2 + \epsilon_t/(b_2 + c_2)) \) for any \( \Gamma \).

To see this, note that

\[
MC \left( q_t^* + \frac{\epsilon_t}{b_2 + c_2} (1 - \Gamma) \left( p_t^* + \frac{\epsilon_t}{c_2} \right) , \theta_t \right) = \\
\sum_{t=1}^{\infty} \left( q_t^* + \frac{\epsilon_t}{b_2 + c_2} - (1 - \Gamma) \left( p_t^* + \frac{\epsilon_t}{b_2 + c_2} \right) \right) = \Gamma \left( p_t^* + \frac{\epsilon_t}{b_2 + c_2} \right)
\]

That is, all the marginal costs are scaled by \( \Gamma \) relative to the original solution, \( p_t^* + c_2 \epsilon_t/(b_2 + c_2) \). Given the original solution satisfied the first-order condition, scaling both sides by \( \Gamma \) will satisfy the condition as well. In particular, note that \( \Gamma = 0 \) will lead to zero marginal cost and a minimum of the cost function in all periods.

Consider, however, what happens to the bank when \( 0 < \Gamma < 1 \). That is, what happens when firms try to reduce costs further relative to the original solution? Under the original solution,

\[
B_t = - \left( q_t^* - E[ q_t^* | t - 1 ] + \frac{\epsilon_t - \epsilon_{t-1}}{b_2 + c_2} \right) = - \frac{\mu_t - \nu_t + \omega_t}{b_2 + c_2}
\]

Note that the original bank is a mean-zero stationary process, with an expectation of zero in advance of each period. Now, consider what happens if the firm (in an effort to lower costs) increases \( q_t \) by \(-(1 - \Gamma)(p_t^*/c_2 + \epsilon_t/(b_2 + c_2))\) each period. With the same government allocation, the bank after \( t \) periods equals

\[
B_t = - \left( q_t^* - E[ q_t^* | t - 1 ] + \frac{\epsilon_t - \epsilon_{t-1}}{b_2 + c_2} \right) + (1 - \Gamma) \sum_{s=1}^{t} \left( \prod_{r=1}^{s} \tilde{R}_r \right) \left( p_t^* + \frac{\epsilon_s}{c_2} + \frac{\epsilon_s}{b_2 + c_2} \right)
\]

Assuming some effort is required by government policy, we have \( p_t^* \neq 0 \). Therefore, unlike the original bank, the new term adds a non-zero expectation. Moreover, this
term grows with \( t \). For example, if the firm takes no meaningful action (so \( p_t = 0 \ \forall t \)), the bank accumulates the emission reductions required to meet the optimal quantity each period \( q_t^* - (\tilde{q}_t - (c_1 + \theta_t)/c_2) \). This will eventually exceed any finite limit on the bank (as shown rigorously in Appendix A of Newell et al. 2005). Therefore, a government policy that includes a finite banking and borrowing limit, coupled with the previous allocation \( \tilde{q}_t \) and trading ratio \( \tilde{R}_t \), will continue to achieve the same outcome with an infinite horizon as the previous model achieved with a finite horizon.

QED.

B.2.5 Comparative Advantage

In the previous sections, we showed that in each period we have a deviation from \( q_t^* \) under the price policy of \( \frac{(\epsilon_{t-1} - (\mu_t + (b_2/c_2)\nu_t))}{b_2 + c_2} \) and a deviation under the quantity policy of \( \frac{\epsilon_t}{b_2 + c_2} \). These correspond to deadweight losses of \( \frac{(\epsilon_{t-1} - (\mu_t + (b_2/c_2)\nu_t))^2}{2(b_2 + c_2)} \) and \( \frac{\epsilon_t^2}{2(b_2 + c_2)} \), respectively. We can compute the expected welfare loss per period for each policy, relative to the first best, as

\[
-(\sigma_{\epsilon_{t-1}}^2 + \sigma_\mu^2 + (b_2/c_2)^2\sigma_\nu^2) \over 2(b_2 + c_2)
\]

for the price policy and

\[
-\frac{\sigma_{\epsilon_t}^2}{2(b_2 + c_2)}
\]

for the quantity policy. We can take the difference to find the comparative advantage of noisy prices compared to noisy quantities,

\[
\Delta^N = \frac{\sigma_\omega^2 - (\sigma_\mu^2 + (b_2/c_2)^2\sigma_\nu^2)}{2(b_2 + c_2)},
\]

where \( \sigma_\omega^2 = \sigma_{\epsilon_t}^2 - \sigma_{\epsilon_{t-1}}^2 \) under the assumption \( \epsilon_t = \epsilon_{t-1} + \omega_t \).

178
Note: This has assumed a perspective evaluating the loss from period 0. Evaluating the expected welfare loss in period $t - 1$, we would have

$$\frac{-\left(\epsilon_{t-1}^2 + \sigma_\mu^2 + (b_2/c_2)^2 \sigma_\nu^2\right)}{2(b_2 + c_2)}$$

as the loss for the price policy and

$$\frac{-\left(\epsilon_{t-1}^2 + \sigma_\omega^2\right)}{2(b_2 + c_2)}$$

for the quantity policy. Again, the difference would be given by the $\Delta^N$ expression above. Note that this also holds even if the government at $t - 1$ does not consider $\epsilon_{t-1}$ to be “noise,” as it cancels between prices and quantities either way.

QED.
Appendix C

Appendix to Chapter 4

C.1 Illustrative Decline Curves

Figure C.1 shows the production profiles for two typical gas wells in our data, both of which began producing in 2007. The unconventional well is located in Johnson County, TX, overlaying the Barnett shale play. The conventional well is located in Shelby County, TX, in the East Texas Basin. This graph, along with Table 4.1, illustrates that while the level-decline in gas production is larger for the unconventional well, the conventional and unconventional wells have quite similar percentage decline rates.
Figure C.1: Production Profile of Typical Gas Wells

Sources: Authors’ calculations based on data from Drillinginfo and EIA

C.2 Gas Well Productivity Over Time

As mentioned in section 4.3.3, unconventional productivity rose substantially during 2005-2015. Figure C.2 shows trends in the productivity of unconventional and conventional gas wells from 2000 to 2015 (2005-2015 for unconventional, as explained in section 4.3.3), measured by the average first full month of production of all wells drilled in the prior two quarters. The figure shows the increase in average unconventional well productivity, which has nearly doubled since 2005. This productivity improvement demonstrates how, for a given price of gas, each unconventional well has generated increasing revenue, and thus how focusing solely on gas prices can

1 We use a two-quarter average to avoid the noise that would be created by focusing only on wells drilled in individual months, some of which involve small numbers of observations, particularly given the small number of conventional wells drilled in recent years as illustrated in Figure 4.4.
misstate changes in the revenues gained from drilling.

**Figure C.2**: Average Gas Production Per Well During the First Full Month, 2000-2015, Quarterly

*Sources:* Authors’ calculations based on data from Drillinginfo and EIA
C.3 Estimation of the Spud-to-Production Time Hazard Models

In this section, we explain how we estimate the spud-to-production time hazard models. Denote the hazard function of well $i$ of type $j$ (conventional or unconventional) and beginning production $t$ months after it was spudded by:

$$h(t, X_{i,j,t}, \theta_j) = \frac{f(t, X_{i,j,t}, \theta_j)}{1 - F(t, X_{i,j,t}, \theta_j)},$$

where $f(t, X_{i,j,t}, \theta_j)$ is the density function of the spud-to-production time, and $F(t, X_{i,j,t}, \theta_j)$ is the corresponding cumulative distribution function. $X_{i,j,t}$ is a vector of explanatory variables, some of which may change over time (for example, gas and oil prices can and do change over time). Hence, the hazard function describes the probability that a well with characteristics $X_{i,j,t}$ will begin production $t$ months after it was spudded, given that it has not yet begun producing.

We estimate this model using maximum likelihood by assuming an underlying density $f(\cdot)$ to specify the form of the baseline hazard function. We focus on using the generalized gamma distribution, which is a very flexible distribution that is parameterized by two ancillary parameters determining the shape of the distribution.\(^2\) We found that the gamma distribution results in an estimated distribution that closely resembles the non-parametric distributions and also yields a significantly higher log-likelihood than other alternatives.\(^3\)

\(^2\) The density of the gamma distribution is given by,

$$f(t) = \begin{cases} \frac{\gamma^\gamma}{\sigma^{\gamma} \Gamma(\gamma)} \exp(z \sqrt{\gamma} - u) & \text{if } \kappa \neq 0 \\ \frac{1}{\sigma \sqrt{2\pi}} \exp(-z^2/2) & \text{if } \kappa = 0, \end{cases}$$

where $\gamma = |\kappa|^{-2}$, $z = \text{sign}(\kappa)(\ln(t) - \mu)/\sigma$, $u = \gamma \exp(|\kappa|z)$, $\Gamma(\cdot)$ is the gamma function, and we parameterize $\mu = X_{i,j,t}\theta_j$. We estimate the ancillary parameters $\sigma$ and $\kappa$ from the data.

\(^3\) We test the Weibull, exponential, Gompertz, Log-normal, and Log-logistic distributions. Several of these distributions are special cases of the gamma distribution (namely, the Weibull, exponential, and log-normal distributions). For those nested distributions we can test for whether the gamma
The shape of the gamma distribution of spud-to-production time is estimated as part of the hazard estimation, separately for each well type. Here, we show evidence that the gamma distribution is a good approximation of a non-parametric estimate of these distributions. The estimated (gamma) distributions are plotted in Figure C.3 at covariate means, along with the non-parametric, kernel density estimates of the underlying distributions. The fitted gamma distributions strongly resemble the non-parametrically estimated densities, suggesting that the gamma distribution fits the true baseline hazard distribution well and is not driving the coefficient estimates. The estimated shape also demonstrates that monotonic functional forms (exponential, Gompertz, Weibull) would be inappropriate.

distribution’s better fit is statistically significant: we find it is in all cases. We also compared the results to a Cox proportional hazards model, which leaves the baseline hazard unspecified: the key estimates were generally in the same directions and of similar relative sizes.
As mentioned in section 4.3.3, in percentage terms, the decline curves are not very different between unconventional and conventional wells in our data. Figure C.4 shows the average decline curves scaled as a percent of peak production. All four curves appear very similar, perhaps with the exception of median conventional wells, suggesting that the main difference between unconventional and conventional gas wells in Texas are in the magnitude, rather than shape, of the production profile.

---

4 Not all curves achieve 100 percent of peak production because these curves represent averages and not all wells achieve their peak production in the same month.
Figure C.4: Mean and Median Profiles of Monthly Gas Production from Gas Wells, as a Percent of Peak Production

Sources: Authors' calculations based on data from Drillinginfo and EIA

C.5 Spud Regression First Stage Estimates
**Table C.1: Revenues Model First Stage (Column 1 of Table 4.2)**

<table>
<thead>
<tr>
<th></th>
<th>Contemp.</th>
<th>1 Lag</th>
<th>2 Lags</th>
<th>Contemp.</th>
<th>1 Lag</th>
<th>2 Lags</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gas Revenue</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDD, Contemporaneous</td>
<td>0.025</td>
<td>-0.117</td>
<td>-0.073</td>
<td>-0.033</td>
<td>-0.103</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.098)</td>
<td>(0.033)</td>
<td>(0.088)</td>
<td>(0.070)</td>
<td>(0.052)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>HDD, 1 Lag</td>
<td>0.031</td>
<td>-0.038</td>
<td>-0.247</td>
<td>0.078</td>
<td>-0.142</td>
<td>-0.155</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.094)</td>
<td>(0.089)</td>
<td>(0.080)</td>
<td>(0.120)</td>
<td>(0.094)</td>
</tr>
<tr>
<td>HDD, 2 Lags</td>
<td>-0.015</td>
<td>0.062</td>
<td>-0.186</td>
<td>0.112</td>
<td>0.017</td>
<td>-0.210</td>
</tr>
<tr>
<td></td>
<td>(0.179)</td>
<td>(0.093)</td>
<td>(0.090)</td>
<td>(0.065)</td>
<td>(0.078)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>HDD, 3 Lags</td>
<td>0.020</td>
<td>0.018</td>
<td>-0.088</td>
<td>0.120</td>
<td>0.033</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td>(0.188)</td>
<td>(0.141)</td>
<td>(0.101)</td>
<td>(0.103)</td>
<td>(0.070)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>HDD, 4 Lags</td>
<td>-0.003</td>
<td>0.137</td>
<td>-0.107</td>
<td>-0.029</td>
<td>0.119</td>
<td>-0.351</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.164)</td>
<td>(0.077)</td>
<td>(0.048)</td>
<td>(0.099)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>CDD, Contemp.</td>
<td>-0.256</td>
<td>0.059</td>
<td>-0.031</td>
<td>0.045</td>
<td>0.014</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.059)</td>
<td>(0.097)</td>
<td>(0.050)</td>
<td>(0.060)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>CDD, 1 Lag</td>
<td>-0.232</td>
<td>-0.049</td>
<td>0.005</td>
<td>0.086</td>
<td>0.037</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.044)</td>
<td>(0.127)</td>
<td>(0.091)</td>
<td>(0.071)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>CDD, 2 Lags</td>
<td>-0.208</td>
<td>0.021</td>
<td>-0.157</td>
<td>0.157</td>
<td>0.062</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(0.136)</td>
<td>(0.073)</td>
<td>(0.106)</td>
<td>(0.103)</td>
<td>(0.073)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>CDD, 3 Lags</td>
<td>-0.070</td>
<td>-0.072</td>
<td>-0.158</td>
<td>0.158</td>
<td>0.049</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.105)</td>
<td>(0.093)</td>
<td>(0.117)</td>
<td>(0.067)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>CDD, 4 Lags</td>
<td>-0.022</td>
<td>0.061</td>
<td>-0.206</td>
<td>0.057</td>
<td>0.037</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.102)</td>
<td>(0.053)</td>
<td>(0.077)</td>
<td>(0.078)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, Contemp.</td>
<td>0.952</td>
<td>-0.429</td>
<td>-1.045</td>
<td>1.430</td>
<td>-1.201</td>
<td>-0.702</td>
</tr>
<tr>
<td></td>
<td>(0.275)</td>
<td>(0.457)</td>
<td>(0.318)</td>
<td>(0.334)</td>
<td>(0.364)</td>
<td>(0.486)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 1 Lag</td>
<td>-0.037</td>
<td>1.456</td>
<td>0.326</td>
<td>0.562</td>
<td>2.193</td>
<td>-0.643</td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>(0.360)</td>
<td>(0.550)</td>
<td>(0.305)</td>
<td>(0.446)</td>
<td>(0.452)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 2 Lags</td>
<td>0.855</td>
<td>-0.320</td>
<td>1.088</td>
<td>0.189</td>
<td>0.363</td>
<td>1.973</td>
</tr>
<tr>
<td></td>
<td>(0.525)</td>
<td>(0.207)</td>
<td>(0.356)</td>
<td>(0.292)</td>
<td>(0.319)</td>
<td>(0.507)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 3 Lags</td>
<td>-0.289</td>
<td>0.438</td>
<td>-0.352</td>
<td>-0.051</td>
<td>-0.008</td>
<td>0.304</td>
</tr>
<tr>
<td></td>
<td>(0.278)</td>
<td>(0.239)</td>
<td>(0.333)</td>
<td>(0.334)</td>
<td>(0.263)</td>
<td>(0.378)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 4 Lags</td>
<td>-0.142</td>
<td>-0.148</td>
<td>0.363</td>
<td>-0.193</td>
<td>-0.078</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.300)</td>
<td>(0.305)</td>
<td>(0.286)</td>
<td>(0.209)</td>
<td>(0.240)</td>
<td>(0.290)</td>
</tr>
<tr>
<td>Gas Inventories, 1 Lag</td>
<td>0.253</td>
<td>-0.578</td>
<td>-0.064</td>
<td>-0.056</td>
<td>-0.217</td>
<td>-0.243</td>
</tr>
<tr>
<td></td>
<td>(0.160)</td>
<td>(0.194)</td>
<td>(0.114)</td>
<td>(0.159)</td>
<td>(0.104)</td>
<td>(0.171)</td>
</tr>
<tr>
<td>Gas Inventories, 2 Lags</td>
<td>0.184</td>
<td>0.132</td>
<td>-0.815</td>
<td>-0.034</td>
<td>-0.087</td>
<td>-0.147</td>
</tr>
<tr>
<td></td>
<td>(0.232)</td>
<td>(0.117)</td>
<td>(0.223)</td>
<td>(0.143)</td>
<td>(0.139)</td>
<td>(0.138)</td>
</tr>
<tr>
<td>Gas Inventories, 3 Lags</td>
<td>0.171</td>
<td>-0.136</td>
<td>0.190</td>
<td>0.189</td>
<td>-0.004</td>
<td>-0.185</td>
</tr>
<tr>
<td></td>
<td>(0.194)</td>
<td>(0.049)</td>
<td>(0.099)</td>
<td>(0.239)</td>
<td>(0.086)</td>
<td>(0.141)</td>
</tr>
<tr>
<td>Gas Inventories, 4 Lags</td>
<td>0.211</td>
<td>0.141</td>
<td>-0.066</td>
<td>0.240</td>
<td>0.283</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>(0.108)</td>
<td>(0.153)</td>
<td>(0.104)</td>
<td>(0.100)</td>
<td>(0.133)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>2nd Quarter Indicator</td>
<td>0.568</td>
<td>-1.349</td>
<td>0.045</td>
<td>-0.480</td>
<td>-0.586</td>
<td>1.041</td>
</tr>
<tr>
<td></td>
<td>(0.611)</td>
<td>(0.628)</td>
<td>(0.498)</td>
<td>(0.547)</td>
<td>(0.481)</td>
<td>(0.413)</td>
</tr>
<tr>
<td>3rd Quarter Indicator</td>
<td>1.052</td>
<td>-1.368</td>
<td>-1.883</td>
<td>-1.055</td>
<td>-1.251</td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td>(1.356)</td>
<td>(0.745)</td>
<td>(0.987)</td>
<td>(0.446)</td>
<td>(0.897)</td>
<td>(0.624)</td>
</tr>
<tr>
<td>4th Quarter Indicator</td>
<td>0.742</td>
<td>-0.460</td>
<td>-1.865</td>
<td>0.285</td>
<td>-1.462</td>
<td>-0.653</td>
</tr>
<tr>
<td></td>
<td>(0.529)</td>
<td>(0.335)</td>
<td>(0.599)</td>
<td>(0.472)</td>
<td>(0.564)</td>
<td>(0.263)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.599</td>
<td>0.790</td>
<td>0.929</td>
<td>0.348</td>
<td>0.866</td>
<td>-0.091</td>
</tr>
<tr>
<td></td>
<td>(0.606)</td>
<td>(0.372)</td>
<td>(0.415)</td>
<td>(0.279)</td>
<td>(0.459)</td>
<td>(0.238)</td>
</tr>
</tbody>
</table>

<p>| Observations | 63 | 63 | 63 | 63 | 63 | 63 |
| R²           | 0.436 | 0.660 | 0.703 | 0.561 | 0.634 | 0.634 |
| Adjusted R²  | 0.126 | 0.474 | 0.320 | 0.430 | 0.433 | 0.433 |
| F Statistic  | 30.4 | 31.1 | 41.1 | 29.6 | 40.2 | 111.0 |</p>
<table>
<thead>
<tr>
<th>Dependent variable</th>
<th></th>
<th>Gas Price</th>
<th></th>
<th>Oil Price</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contemp.</td>
<td>1 Lag</td>
<td>2 Lags</td>
<td>Contemp.</td>
<td>1 Lag</td>
</tr>
<tr>
<td>HDD, Contemporaneous</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>−0.006</td>
<td>−0.125</td>
<td>−0.028</td>
<td>−0.036</td>
<td>−0.073</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.046)</td>
<td>(0.081)</td>
<td>(0.060)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>HDD, 1 Lag</td>
<td>0.0001</td>
<td>−0.064</td>
<td>−0.194</td>
<td>0.079</td>
<td>−0.114</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.080)</td>
<td>(0.084)</td>
<td>(0.0619)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>HDD, 2 Lags</td>
<td>−0.090</td>
<td>0.035</td>
<td>−0.146</td>
<td>0.051</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td>(0.080)</td>
<td>(0.084)</td>
<td>(0.053)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>HDD, 3 Lags</td>
<td>−0.043</td>
<td>−0.044</td>
<td>−0.051</td>
<td>0.077</td>
<td>−0.039</td>
</tr>
<tr>
<td></td>
<td>(0.162)</td>
<td>(0.136)</td>
<td>(0.074)</td>
<td>(0.058)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>HDD, 4 Lags</td>
<td>−0.023</td>
<td>0.086</td>
<td>−0.117</td>
<td>−0.005</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.152)</td>
<td>(0.064)</td>
<td>(0.044)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>CDD, Contemp.</td>
<td>−0.260</td>
<td>0.045</td>
<td>−0.004</td>
<td>0.024</td>
<td>−0.029</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.080)</td>
<td>(0.087)</td>
<td>(0.035)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>CDD, 1 Lag</td>
<td>−0.243</td>
<td>−0.053</td>
<td>0.025</td>
<td>0.047</td>
<td>−0.027</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.058)</td>
<td>(0.110)</td>
<td>(0.087)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>CDD, 2 Lags</td>
<td>−0.286</td>
<td>0.018</td>
<td>−0.111</td>
<td>0.084</td>
<td>−0.040</td>
</tr>
<tr>
<td></td>
<td>(0.174)</td>
<td>(0.074)</td>
<td>(0.091)</td>
<td>(0.099)</td>
<td>(0.092)</td>
</tr>
<tr>
<td>CDD, 3 Lags</td>
<td>−0.119</td>
<td>−0.118</td>
<td>−0.101</td>
<td>0.128</td>
<td>−0.045</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.102)</td>
<td>(0.084)</td>
<td>(0.090)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>CDD, 4 Lags</td>
<td>−0.039</td>
<td>0.044</td>
<td>−0.222</td>
<td>0.081</td>
<td>−0.041</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.092)</td>
<td>(0.047)</td>
<td>(0.061)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, Contemp.</td>
<td>0.836</td>
<td>−0.290</td>
<td>−0.940</td>
<td>1.383</td>
<td>−0.884</td>
</tr>
<tr>
<td></td>
<td>(0.465)</td>
<td>(0.465)</td>
<td>(0.223)</td>
<td>(0.207)</td>
<td>(0.284)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 1 Lag</td>
<td>−0.105</td>
<td>1.262</td>
<td>0.338</td>
<td>0.323</td>
<td>1.886</td>
</tr>
<tr>
<td></td>
<td>(0.324)</td>
<td>(0.406)</td>
<td>(0.515)</td>
<td>(0.179)</td>
<td>(0.268)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 2 Lags</td>
<td>0.747</td>
<td>−0.385</td>
<td>0.937</td>
<td>0.072</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>(0.511)</td>
<td>(0.213)</td>
<td>(0.285)</td>
<td>(0.194)</td>
<td>(0.177)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 3 Lags</td>
<td>−0.385</td>
<td>0.267</td>
<td>−0.360</td>
<td>−0.147</td>
<td>−0.085</td>
</tr>
<tr>
<td></td>
<td>(0.350)</td>
<td>(0.225)</td>
<td>(0.278)</td>
<td>(0.201)</td>
<td>(0.188)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 4 Lags</td>
<td>0.027</td>
<td>−0.138</td>
<td>0.202</td>
<td>0.217</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>(0.374)</td>
<td>(0.277)</td>
<td>(0.273)</td>
<td>(0.189)</td>
<td>(0.276)</td>
</tr>
<tr>
<td>Gas Inventories, 1 Lag</td>
<td>0.214</td>
<td>−0.863</td>
<td>−0.033</td>
<td>−0.104</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.214)</td>
<td>(0.180)</td>
<td>(0.124)</td>
<td>(0.172)</td>
<td>(0.123)</td>
</tr>
<tr>
<td>Gas Inventories, 2 Lags</td>
<td>0.140</td>
<td>0.070</td>
<td>−0.809</td>
<td>−0.137</td>
<td>−0.068</td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.120)</td>
<td>(0.173)</td>
<td>(0.101)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Gas Inventories, 3 Lags</td>
<td>0.132</td>
<td>−0.179</td>
<td>0.158</td>
<td>0.079</td>
<td>−0.030</td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.075)</td>
<td>(0.074)</td>
<td>(0.182)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Gas Inventories, 4 Lags</td>
<td>0.136</td>
<td>0.061</td>
<td>−0.110</td>
<td>0.046</td>
<td>0.143</td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td>(0.128)</td>
<td>(0.084)</td>
<td>(0.086)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>2nd Quarter Indicator</td>
<td>0.572</td>
<td>−1.368</td>
<td>0.224</td>
<td>−0.239</td>
<td>−0.548</td>
</tr>
<tr>
<td></td>
<td>(0.754)</td>
<td>(0.690)</td>
<td>(0.406)</td>
<td>(0.498)</td>
<td>(0.719)</td>
</tr>
<tr>
<td>3rd Quarter Indicator</td>
<td>1.044</td>
<td>−1.345</td>
<td>−1.500</td>
<td>−0.469</td>
<td>−0.749</td>
</tr>
<tr>
<td></td>
<td>(1.519)</td>
<td>(0.768)</td>
<td>(0.882)</td>
<td>(0.626)</td>
<td>(0.820)</td>
</tr>
<tr>
<td>4th Quarter Indicator</td>
<td>0.865</td>
<td>−0.420</td>
<td>−1.596</td>
<td>0.625</td>
<td>−0.906</td>
</tr>
<tr>
<td></td>
<td>(0.567)</td>
<td>(0.505)</td>
<td>(0.402)</td>
<td>(0.560)</td>
<td>(0.480)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.634</td>
<td>0.775</td>
<td>0.715</td>
<td>0.015</td>
<td>0.550</td>
</tr>
<tr>
<td></td>
<td>(0.676)</td>
<td>(0.405)</td>
<td>(0.351)</td>
<td>(0.365)</td>
<td>(0.432)</td>
</tr>
<tr>
<td>Observations</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>R^2</td>
<td>0.383</td>
<td>0.642</td>
<td>0.695</td>
<td>0.635</td>
<td>0.695</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.043</td>
<td>0.445</td>
<td>0.527</td>
<td>0.435</td>
<td>0.527</td>
</tr>
<tr>
<td>F Statistic</td>
<td>11.2</td>
<td>30.7</td>
<td>39.8</td>
<td>46.4</td>
<td>39.6</td>
</tr>
</tbody>
</table>
Table C.3: Prices Plus Productivity First Stage (Column 3 of Table 4.2)

<table>
<thead>
<tr>
<th></th>
<th>Gas Price</th>
<th>Oil Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contemp.</td>
<td>1 Lag</td>
</tr>
<tr>
<td>HDD, Contemporaneous</td>
<td>−0.006</td>
<td>−0.125</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>HDD, 1 Lag</td>
<td>0.0001</td>
<td>−0.064</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>HDD, 2 Lags</td>
<td>−0.090</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>HDD, 3 Lags</td>
<td>−0.043</td>
<td>−0.044</td>
</tr>
<tr>
<td></td>
<td>(0.162)</td>
<td>(0.136)</td>
</tr>
<tr>
<td>HDD, 4 Lags</td>
<td>−0.023</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.152)</td>
</tr>
<tr>
<td>CDD, Contemp.</td>
<td>−0.260</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>CDD, 1 Lag</td>
<td>−0.243</td>
<td>−0.053</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>CDD, 2 Lags</td>
<td>−0.286</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.174)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>CDD, 3 Lags</td>
<td>−0.119</td>
<td>−0.118</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>CDD, 4 Lags</td>
<td>−0.039</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.092)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, Contemp.</td>
<td>0.836</td>
<td>−0.290</td>
</tr>
<tr>
<td></td>
<td>(0.465)</td>
<td>(0.465)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 1 Lag</td>
<td>−0.105</td>
<td>1.262</td>
</tr>
<tr>
<td></td>
<td>(0.324)</td>
<td>(0.406)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 2 Lags</td>
<td>0.747</td>
<td>−0.385</td>
</tr>
<tr>
<td></td>
<td>(0.511)</td>
<td>(0.213)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 3 Lags</td>
<td>−0.385</td>
<td>0.267</td>
</tr>
<tr>
<td></td>
<td>(0.350)</td>
<td>(0.225)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 4 Lags</td>
<td>0.027</td>
<td>−0.138</td>
</tr>
<tr>
<td></td>
<td>(0.374)</td>
<td>(0.277)</td>
</tr>
<tr>
<td>Gas Inventories, 1 Lag</td>
<td>0.214</td>
<td>−0.863</td>
</tr>
<tr>
<td></td>
<td>(0.214)</td>
<td>(0.180)</td>
</tr>
<tr>
<td>Gas Inventories, 2 Lags</td>
<td>0.140</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.120)</td>
</tr>
<tr>
<td>Gas Inventories, 3 Lags</td>
<td>0.132</td>
<td>−0.179</td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>Gas Inventories, 4 Lags</td>
<td>0.136</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td>(0.128)</td>
</tr>
<tr>
<td>2nd Quarter Indicator</td>
<td>0.572</td>
<td>−1.368</td>
</tr>
<tr>
<td></td>
<td>(0.754)</td>
<td>(0.690)</td>
</tr>
<tr>
<td>3rd Quarter Indicator</td>
<td>1.044</td>
<td>−1.345</td>
</tr>
<tr>
<td></td>
<td>(1.519)</td>
<td>(0.768)</td>
</tr>
<tr>
<td>4th Quarter Indicator</td>
<td>0.865</td>
<td>−0.420</td>
</tr>
<tr>
<td></td>
<td>(0.567)</td>
<td>(0.505)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.634</td>
<td>0.775</td>
</tr>
<tr>
<td></td>
<td>(0.676)</td>
<td>(0.405)</td>
</tr>
</tbody>
</table>

Observations 63 63 63 63 63 63

R² 0.383 0.642 0.695 0.635 0.695 0.698

Adjusted R² 0.043 0.445 0.527 0.435 0.527 0.532

F Statistic 11.2 30.7 39.8 46.4 39.6 210.4

189
### Table C.4: Unconventional-Only First Stage (Column 4 of Table 4.2)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Gas Revenue</th>
<th>Oil Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contemp.</td>
<td>1 Lag</td>
</tr>
<tr>
<td>HDD, Contemporaneous</td>
<td>-0.148</td>
<td>-0.165</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>HDD, 1 Lag</td>
<td>-0.191</td>
<td>-0.297</td>
</tr>
<tr>
<td></td>
<td>(0.185)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>HDD, 2 Lags</td>
<td>-0.335</td>
<td>-0.070</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.154)</td>
</tr>
<tr>
<td>HDD, 3 Lags</td>
<td>-0.195</td>
<td>-0.214</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.147)</td>
</tr>
<tr>
<td>HDD, 4 Lags</td>
<td>-0.014</td>
<td>-0.081</td>
</tr>
<tr>
<td></td>
<td>(0.151)</td>
<td>(0.127)</td>
</tr>
<tr>
<td>CDD, Contemp.</td>
<td>-0.278</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>CDD, 1 Lag</td>
<td>-0.197</td>
<td>-0.107</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>CDD, 2 Lags</td>
<td>-0.251</td>
<td>-0.044</td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>CDD, 3 Lags</td>
<td>-0.106</td>
<td>-0.214</td>
</tr>
<tr>
<td></td>
<td>(0.171)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>CDD, 4 Lags</td>
<td>-0.116</td>
<td>-0.045</td>
</tr>
<tr>
<td></td>
<td>(0.140)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, Contemp.</td>
<td>0.790</td>
<td>-0.800</td>
</tr>
<tr>
<td></td>
<td>(0.458)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 1 Lag</td>
<td>-0.180</td>
<td>1.684</td>
</tr>
<tr>
<td></td>
<td>(0.444)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 2 Lags</td>
<td>1.267</td>
<td>-0.433</td>
</tr>
<tr>
<td></td>
<td>(0.401)</td>
<td>(0.278)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 3 Lags</td>
<td>-0.946</td>
<td>0.450</td>
</tr>
<tr>
<td></td>
<td>(0.769)</td>
<td>(0.437)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 4 Lags</td>
<td>-0.587</td>
<td>-0.705</td>
</tr>
<tr>
<td></td>
<td>(0.304)</td>
<td>(0.164)</td>
</tr>
<tr>
<td>Gas Inventories, 1 Lag</td>
<td>-0.107</td>
<td>-0.734</td>
</tr>
<tr>
<td></td>
<td>(0.339)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Gas Inventories, 2 Lags</td>
<td>-0.197</td>
<td>-0.111</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.157)</td>
</tr>
<tr>
<td>Gas Inventories, 3 Lags</td>
<td>-0.112</td>
<td>-0.118</td>
</tr>
<tr>
<td></td>
<td>(0.192)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>Gas Inventories, 4 Lags</td>
<td>0.049</td>
<td>-0.149</td>
</tr>
<tr>
<td></td>
<td>(0.183)</td>
<td>(0.180)</td>
</tr>
<tr>
<td>2nd Quarter Indicator</td>
<td>1.531</td>
<td>-2.216</td>
</tr>
<tr>
<td></td>
<td>(0.668)</td>
<td>(0.458)</td>
</tr>
<tr>
<td>3rd Quarter Indicator</td>
<td>1.526</td>
<td>-2.514</td>
</tr>
<tr>
<td></td>
<td>(2.021)</td>
<td>(1.216)</td>
</tr>
<tr>
<td>4th Quarter Indicator</td>
<td>0.422</td>
<td>-1.179</td>
</tr>
<tr>
<td></td>
<td>(0.699)</td>
<td>(0.714)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.876</td>
<td>1.473</td>
</tr>
<tr>
<td></td>
<td>(0.766)</td>
<td>(0.565)</td>
</tr>
</tbody>
</table>

| Observations | 43 | 43 | 43 | 43 | 43 |
| R² | 0.700 | 0.829 | 0.829 | 0.598 | 0.684 | 0.654 |
| Adjusted R² | 0.370 | 0.641 | 0.641 | 0.156 | 0.336 | 0.273 |
| F Statistic | 23.2 | 152.3 | 191.1 | 22.2 | 31.2 | 43.1 |

Note: Revenues are computed for unconventional wells, using unconventional productivity estimates.
<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Contemp.</th>
<th>1 Lag</th>
<th>2 Lags</th>
<th>Contemp.</th>
<th>1 Lag</th>
<th>2 Lags</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDD, Contemporaneous</td>
<td>0.055</td>
<td>−0.100</td>
<td>−0.058</td>
<td>0.101</td>
<td>0.073</td>
<td>−0.043</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.052)</td>
<td>(0.095)</td>
<td>(0.124)</td>
<td>(0.067)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>HDD, 1 Lag</td>
<td>0.085</td>
<td>−0.003</td>
<td>−0.235</td>
<td>0.190</td>
<td>0.155</td>
<td>−0.049</td>
</tr>
<tr>
<td></td>
<td>(0.188)</td>
<td>(0.101)</td>
<td>(0.114)</td>
<td>(0.171)</td>
<td>(0.105)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>HDD, 2 Lags</td>
<td>0.038</td>
<td>0.086</td>
<td>−0.193</td>
<td>0.222</td>
<td>0.264</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.132)</td>
<td>(0.130)</td>
<td>(0.167)</td>
<td>(0.087)</td>
<td>(0.105)</td>
</tr>
<tr>
<td>HDD, 3 Lags</td>
<td>0.098</td>
<td>0.044</td>
<td>−0.085</td>
<td>0.171</td>
<td>0.228</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>(0.246)</td>
<td>(0.164)</td>
<td>(0.146)</td>
<td>(0.175)</td>
<td>(0.120)</td>
<td>(0.130)</td>
</tr>
<tr>
<td>HDD, 4 Lags</td>
<td>0.059</td>
<td>0.181</td>
<td>−0.098</td>
<td>−0.029</td>
<td>0.220</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>(0.162)</td>
<td>(0.207)</td>
<td>(0.098)</td>
<td>(0.143)</td>
<td>(0.126)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>CDD, Contemp.</td>
<td>−0.266</td>
<td>0.005</td>
<td>−0.113</td>
<td>0.093</td>
<td>0.046</td>
<td>−0.059</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.075)</td>
<td>(0.124)</td>
<td>(0.065)</td>
<td>(0.127)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>CDD, 1 Lag</td>
<td>−0.192</td>
<td>−0.076</td>
<td>−0.115</td>
<td>0.095</td>
<td>0.185</td>
<td>−0.001</td>
</tr>
<tr>
<td></td>
<td>(0.165)</td>
<td>(0.090)</td>
<td>(0.153)</td>
<td>(0.171)</td>
<td>(0.095)</td>
<td>(0.180)</td>
</tr>
<tr>
<td>CDD, 2 Lags</td>
<td>−0.130</td>
<td>0.042</td>
<td>−0.251</td>
<td>0.172</td>
<td>0.229</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>(0.192)</td>
<td>(0.117)</td>
<td>(0.134)</td>
<td>(0.203)</td>
<td>(0.142)</td>
<td>(0.164)</td>
</tr>
<tr>
<td>CDD, 3 Lags</td>
<td>−0.020</td>
<td>−0.024</td>
<td>−0.211</td>
<td>0.113</td>
<td>0.192</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.174)</td>
<td>(0.128)</td>
<td>(0.126)</td>
<td>(0.223)</td>
<td>(0.171)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>CDD, 4 Lags</td>
<td>0.023</td>
<td>0.112</td>
<td>−0.210</td>
<td>0.008</td>
<td>0.133</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.131)</td>
<td>(0.067)</td>
<td>(0.118)</td>
<td>(0.157)</td>
<td>(0.120)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, Contemp.</td>
<td>0.817</td>
<td>−0.192</td>
<td>−1.213</td>
<td>1.116</td>
<td>−0.564</td>
<td>−0.433</td>
</tr>
<tr>
<td></td>
<td>(0.365)</td>
<td>(0.520)</td>
<td>(0.427)</td>
<td>(0.509)</td>
<td>(0.726)</td>
<td>(0.820)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 1 Lag</td>
<td>−0.095</td>
<td>1.081</td>
<td>0.572</td>
<td>0.563</td>
<td>1.314</td>
<td>−0.414</td>
</tr>
<tr>
<td></td>
<td>(0.305)</td>
<td>(0.423)</td>
<td>(0.393)</td>
<td>(0.347)</td>
<td>(0.622)</td>
<td>(0.640)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 2 Lags</td>
<td>1.087</td>
<td>−0.250</td>
<td>0.733</td>
<td>0.685</td>
<td>0.791</td>
<td>1.314</td>
</tr>
<tr>
<td></td>
<td>(0.460)</td>
<td>(0.224)</td>
<td>(0.475)</td>
<td>(0.619)</td>
<td>(0.438)</td>
<td>(0.633)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 3 Lags</td>
<td>−0.643</td>
<td>0.428</td>
<td>−0.369</td>
<td>−1.154</td>
<td>0.083</td>
<td>0.606</td>
</tr>
<tr>
<td></td>
<td>(0.265)</td>
<td>(0.279)</td>
<td>(0.328)</td>
<td>(0.403)</td>
<td>(0.390)</td>
<td>(0.475)</td>
</tr>
<tr>
<td>CRB Raw Ind. Index, 4 Lags</td>
<td>0.141</td>
<td>−0.244</td>
<td>0.384</td>
<td>0.394</td>
<td>−0.685</td>
<td>−0.179</td>
</tr>
<tr>
<td></td>
<td>(0.294)</td>
<td>(0.286)</td>
<td>(0.362)</td>
<td>(0.385)</td>
<td>(0.388)</td>
<td>(0.443)</td>
</tr>
<tr>
<td>Gas Inventories, 1 Lag</td>
<td>0.096</td>
<td>−0.971</td>
<td>−0.078</td>
<td>−0.341</td>
<td>−0.487</td>
<td>−0.357</td>
</tr>
<tr>
<td></td>
<td>(0.188)</td>
<td>(0.272)</td>
<td>(0.145)</td>
<td>(0.253)</td>
<td>(0.227)</td>
<td>(0.315)</td>
</tr>
<tr>
<td>Gas Inventories, 2 Lags</td>
<td>0.128</td>
<td>0.041</td>
<td>−0.859</td>
<td>−0.075</td>
<td>−0.307</td>
<td>−0.364</td>
</tr>
<tr>
<td></td>
<td>(0.271)</td>
<td>(0.130)</td>
<td>(0.256)</td>
<td>(0.152)</td>
<td>(0.237)</td>
<td>(0.364)</td>
</tr>
<tr>
<td>Gas Inventories, 3 Lags</td>
<td>0.159</td>
<td>−0.270</td>
<td>0.122</td>
<td>0.243</td>
<td>−0.329</td>
<td>−0.429</td>
</tr>
<tr>
<td></td>
<td>(0.221)</td>
<td>(0.095)</td>
<td>(0.130)</td>
<td>(0.271)</td>
<td>(0.231)</td>
<td>(0.267)</td>
</tr>
<tr>
<td>Gas Inventories, 4 Lags</td>
<td>0.208</td>
<td>0.139</td>
<td>−0.151</td>
<td>0.494</td>
<td>0.296</td>
<td>−0.129</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.166)</td>
<td>(0.170)</td>
<td>(0.160)</td>
<td>(0.192)</td>
<td>(0.337)</td>
</tr>
<tr>
<td>2nd Quarter Indicator</td>
<td>0.775</td>
<td>−1.180</td>
<td>0.165</td>
<td>−0.824</td>
<td>−0.303</td>
<td>0.617</td>
</tr>
<tr>
<td></td>
<td>(0.894)</td>
<td>(0.947)</td>
<td>(0.600)</td>
<td>(0.848)</td>
<td>(0.939)</td>
<td>(0.751)</td>
</tr>
<tr>
<td>3rd Quarter Indicator</td>
<td>1.247</td>
<td>−0.878</td>
<td>−1.462</td>
<td>−1.323</td>
<td>−0.612</td>
<td>−0.286</td>
</tr>
<tr>
<td></td>
<td>(1.952)</td>
<td>(0.945)</td>
<td>(1.247)</td>
<td>(0.886)</td>
<td>(1.672)</td>
<td>(1.272)</td>
</tr>
<tr>
<td>4th Quarter Indicator</td>
<td>0.691</td>
<td>−0.247</td>
<td>−1.575</td>
<td>0.010</td>
<td>−1.208</td>
<td>−0.701</td>
</tr>
<tr>
<td></td>
<td>(0.949)</td>
<td>(0.401)</td>
<td>(0.672)</td>
<td>(0.618)</td>
<td>(0.847)</td>
<td>(0.715)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.706</td>
<td>0.557</td>
<td>0.708</td>
<td>0.539</td>
<td>0.546</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>(0.924)</td>
<td>(0.448)</td>
<td>(0.528)</td>
<td>(0.395)</td>
<td>(0.812)</td>
<td>(0.494)</td>
</tr>
</tbody>
</table>

| Observations | 63     | 63     | 63     | 63     | 63     | 63     |
| R²           | 0.346  | 0.546  | 0.603  | 0.504  | 0.488  | 0.456  |
| Adjusted R²  | −0.013 | 0.296  | 0.385  | 0.231  | 0.207  | 0.158  |
| F Statistic  | 31.8   | 52.2   | 28.4   | 19.9   | 16.8   | 62.3   |

Note: Revenues are computed for conventional wells, using conventional productivity estimates.


James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013), *An Introduction to Statistical Learning*, vol. 6, Springer.


197


Biography

Brian Charles Prest was born in Beverly, Massachusetts on May 9th, 1986. He obtained his degrees in Bachelor of Arts in 2009 from Williams College, Master of Arts in 2014 from Duke University, and Doctor of Philosophy in May 2018 from Duke University.

As of this writing, his article “Trophy Hunting vs. Manufacturing Energy: The Price-Responsiveness of Shale Gas” (also chapter 3 of this dissertation) is forthcoming at the Journal of the Association of Environmental and Resource Economists (Newell et al. 2019).

At Duke, he has been awarded the following honors (awarding institution in parentheses):

• Student Paper Award, (United States Association for Energy Economics)
• Joseph L. Fisher Doctoral Dissertation Fellowship (Resources for the Future)
• Energy Doctoral Student Fellowship (Duke University Energy Initiative)
• Pathfinder Fellow, Data Expedition Grant (Information Initiative at Duke)
• Duke Environmental Economics Doctoral Scholars Scholar (Nicholas Institute for Environmental Policy Solutions)

After completing his doctorate in 2018, he will be joining Resources for the Future as a postdoctoral fellow.