ABSTRACT

Title of Dissertation: ESSAYS IN NONMARKET VALUATION AND ENERGY ECONOMICS

Mark Vincent Miller, Doctor of Philosophy, 2016

Dissertation directed by: Professor Anna Alberini
Department of Agricultural and Resource Economics

This dissertation comprises three essays, relating to negative externalities in economics. The first essay concentrates on residential electricity consumption. In the economic literature, price elasticity of demands estimates for this market vary widely. In this essay, I seek to explain these findings using three nationwide datasets – the American Housing Survey, Forms EIA-861, and the Residential Energy Consumption Survey – from the U.S. I examine the role of the sample period, level of aggregation, use of panel data, use of instrumental variables, and inclusion of housing characteristics and capital stock. The findings suggest that price elasticities have remained relatively constant over the time period considered, from 1997 to 2009. Upon splitting my panel datasets into annual cross sections, I do observe a negative relationship between price elasticities and the average price. Whether prices are rising or falling appears to have little effect on the estimates. I also find that aggregating our data can result in both higher and lower price elasticity estimates, depending on the dataset used, and that controlling for unit-level fixed effects with panel data generally results in more inelastic demand functions. Addressing the endogeneity of price and/or measurement error in price with instrumental

variables has a small but noticeable effect on the price elasticities. Finally, controlling for housing characteristics and capital stock produces a lower price elasticity.

My second essay focuses on personal vehicular transportation, which is the source of several externalities, including congestion and conventional air pollutant and greenhouse gas emissions. In this work, I examine the geographical distribution of carbon dioxide emissions from cars. Specifically, I focus on rural and urban households in the UK, using repeated cross sections from the UK National Travel Survey. I contrast driving behavior with new vehicle purchases, and ask three related questions: First, do rural households purchase more vehicles, and/or vehicles with higher emissions rates? Second, do they drive more miles than their urban counterparts? Third, how do the "carbon footprints" from these two groups compare? To answer these questions, I first model the number of vehicles chosen by each household, and the emissions type of each. I then model the travel demand for that vehicle, conditioning on the latter choice. I use these results to estimate the contribution of rural and urban residents to total CO_2 emissions. I find that rural households own more vehicles than urban households, and that these vehicles have higher emissions rates. Rural vehicles are also driven 12.9% more than urban vehicles. Lastly, I estimate a carbon footprint that is 58% higher for rural households than urban.

My third essay considers an application of the hedonic price models, which are widely used in nonmarket valuation to find the value of environmental quality, changes in health risks, etc. The approach relies on property values. It is less commonly employed, however, to assess the value of certain amenities like crime mitigation. One problem with deploying hedonics in this area is that crime tends to be correlated with unobserved neighborhood attributes, and for that reason regression coefficients on crime may be biased. This chapter considers one particular set of crimes, methamphetamine laboratories, on property values. While methamphetamine labs are more likely to be established and discovered in some neighborhoods more than others, I follow previous literature by arguing that, within the confines of the neighborhoods they are discovered in, they are as good as randomly assigned and can thus be regarded as locally exogenous in regression analysis. I use three registers – one federal, two local – to identify the locations of past discoveries, in Summit County, OH, and categorize the discoveries by scale (e.g., a small discovery of meth in a dumpster versus an actual meth lab). I find mixed evidence that the value of homes closest to a discovery are negatively affected, either directly or in terms of relatively higher properties for properties that are somewhat further away from such discoveries. The effect, when identified, is slightly stronger for large-scale discoveries. I also consider the effect of information disclosure, and find that, for repeat sales observations, property sales that occur following a discovery within 200 meters *and* after public registers are available observe a sizable loss in value. The robustness each of these findings, however, appears questionable, since their magnitude and statistical significance are sensitive to model specification.

ESSAYS IN NONMARKET VALUATION AND ENERGY ECONOMICS

By

Mark Vincent Miller

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor Anna Alberini, Chair
Professor Ted McConnell
Professor Charles Towe
Professor Sebastien Houde
Professor Maureen Cropper

DEDICATION

To my mother and father

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support and help of many people. I am especially grateful for the mentorship, advising, and guidance offered to me by my advisor, Anna, who has worked patiently and extensively with me, over the years. Professor McConnell also provided extensive mentorship, beginning as early as my first year, when he served as my advisor for the qualifying paper. Charles Towe provided me with extensive feedback, from our work together over several years on different projects. I also thank Maureen Cropper and Sebastien Houde for serving on my committee.

Special mentions go to several of my AREC colleagues, whose feedback, compassion, and support are worthy of recognition. Kota Minegishi, Daniel Voica, Ian Page, Elina Tselepidakis, Davide Cerutti, Dan Werner, Jen He, Rubaba Ali, and Shirley Pon offered extensive camaraderie over the years, and offered many opportunities to discuss my work. I owe additional thanks to friends outside the department, who provided unwavering and heartfelt support and feedback, over the years: David Ezell, Sarah Combs, Cat and John Healy, Olivier Francois, Julie Herrick, Charlene Chi-Johnston, and Kumiko Aizawa.

I am also indebted to the many teachers I had prior to entering this program, whose dedication and passion helped shape and stimulate the intellectual curiosity the brought me to, and got me through, this program. This includes Bob Gitter, Goran Skosples, and Julide Yazar from the Ohio Wesleyan University Department of Economics, my original sources of inspiration for getting a Ph.D. in the subject.

Finally, this journey would never have been possible without the perpetual support of my family who saw things through with me, through thick and thin. I am grateful for the patience and sympathy of my brothers, Peter and Christopher, my sister in law Danielle, and my sisters Gabrielle and Jennifer. My mother, Margaret, proved to be the ultimate beacon for unwavering support and encouragement. She was unconditional in her regard for my abilities and talents, and in her support for me pursuing my interests and ambitions, no matter what the costs. My father, James, too, was exceptional in the support he provided to help me see though this chapter in my life. He has been one of the most exemplary figures I've had to learn from. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

Externalities are activities where the actions of one agent affect the utility of another agent, who neither receives nor provides payments for such. As pollution is a negative externality, it is a major topic of study in environmental and energy economics. Pollution externalities are relevant in the context of policymaking, since there are often benefits to be gained from mitigating their production, but no markets in place to facilitate such exchanges.

This dissertation comprises three essays, each of which focuses on a topic relating to the problem of environmental externalities. The first essay concentrates on the market for consumer electricity, in the United States (US). Sixty seven percent of US consumers rely on power plants using fossil fuels, 58% of which being coal, to generate their electricity (US EIA, 2015). The externalities arising from coal-fired power plants, namely the emissions of chemical compounds such as sulfur-dioxide (SO_2), particulate emissions, and carbon dioxide (CO_2) are well recognized, and have been the subject of such major regulatory actions as the Clean Air Act Amendments of 1970, 1977, and 1990. While technological innovations since then have proven effective in reducing this air pollution, such as scrubbers in the case of SO_2, progress remains to be made for others – most notably, CO_2, which remains one of the top pollutants of concern in the current political climate. Carbon taxes have been broached as one possible instrument for reducing these emissions. They would reduce consumption of energy from fossil fuels (and encourage a shift to renewable energy).

How might we expect consumers to respond to a carbon tax? This requires knowledge about the consumers' price elasticity of demand for energy. In the case of electricity, it is generally assumed that consumers are relatively price inelastic, in which case we would expect that a carbon tax would need to be larger for a notable impact. But are consumers necessarily price inelastic? A closer inspection of the economic literature –which includes dozens of studies – suggests that price elasticities for consumer electricity vary widely.[1] What might explain these discrepancies? In my first essay, I seek to explain several of these findings using three nationwide datasets – the American Housing Survey (AHS), Forms EIA-861, and the Residential Energy Consumption Survey (RECS) – from the US. I examine the role of the sample period, level of aggregation, use of panel data, use of instrumental variables, and inclusion of housing characteristics and capital stock.

My findings suggest that price elasticities have remained relatively constant over time. Upon splitting my panel datasets into annual cross sections, I do observe a negative relationship between price elasticities and the average price. Whether prices are rising or falling appears to have little effect on my estimates. I also find that aggregating our data can result in both higher and lower price elasticity estimates, depending on the dataset used, and that controlling for unit-level fixed effects with panel data generally results in more inelastic demand functions. Addressing the endogeneity of price and/or measurement error in price with instrumental variables has a small but noticeable effect on the price elasticities. Finally, I find that controlling for housing characteristics and capital stock produces a lower price elasticity.

[1] The US Energy Information Administration (EIA, 2014), for example, assumes price elasticities for consumer electricity which range from -0.12 to -0.4.

My second essay explores household transportation behavior in the United Kingdom (UK), and its relationship with fuel inputs and CO_2 emissions. As transportation also involves the combustion of fossil fuels, it shares many of the airborne externalities previously noted for electricity production. It is also the source of additional externalities, however, including traffic accidents, congestion, and wear and tear on road infrastructure. In this essay, I focus specifically on the externalities generated from vehicular CO_2 emissions, on their geographical distribution, and on how they compare between urban and rural households. This is done using repeated cross sections from the UK National Travel Survey. I focus on driving behavior associated with new vehicles, and ask three related questions: First, do rural households purchase more vehicles, and/or vehicles with higher emissions rates? Second, do they drive more miles than their urban counterparts? Third, how do the "carbon footprints" from these two groups compare?

To answer these questions, I first model the number of vehicles chosen by each household, and the emissions type of each. I then model the travel demand for that vehicle, conditioning on the latter choice. I use these results to estimate the contribution of rural and urban residents to total CO_2 emissions. I find that rural households own more vehicles than urban households, and that these vehicles have higher emissions rates; however, the difference for the latter is marginal. Rural vehicles are also driven 12.9% more than urban vehicles. Lastly, I use these results to compute a carbon footprint for each group, finding that rural households produce 58% more emissions from vehicles than urban households.

The third chapter of my dissertation considers negative externalities generated by methamphetamine laboratory discoveries, as evidenced by their effect on neighboring property values. Methamphetamines are an illicit substance consumed for recreational drug use, based on a simple production process requiring few inputs. Over the past decade, there has been a surge in small scale operations, which are typically established in residential homes, and often in neighborhoods where they are unexpected. Their presence often goes unrecognized by the community, prior to their discovery by law enforcement, or new residents of a home. Once this occurs, their locations and dates of discovery are publicly disclosed on a register.

There are two reasons to expect a discovery to have a negative effect on property values. First, they are a crime scene, whose presence on a public register could stigmatize a neighborhood by suggesting the potential presence of other criminal activity. Second, they are hazardous waste sites, which could pose a risk to neighboring homes in the area. In this chapter, I use a hedonic price model estimate the effect of these discoveries on property values. Typically, it is difficult to employ this model to the value of disamenities like crime, since they tend to be correlated with other unobserved neighborhood attributes affecting property values, and for that reason regression coefficients on crime may be biased. While methamphetamine labs are more likely to be established and discovered in some neighborhoods more than others, I follow previous literature by arguing that, within the confines of the neighborhoods they are discovered in, they are as good as randomly assigned and can thus be regarded as locally exogenous in regression analysis. I use three registers – one federal, two local – to identify the locations of past discoveries, in Summit County, Ohio, and categorize the discoveries by scale (e.g., a small discovery of meth in a dumpster versus an actual meth lab). I find mixed evidence that the value of homes closest to a discovery are negatively affected, either directly or in terms of relatively higher properties for properties that are somewhat further away from such discoveries. The effect, when identified, is slightly stronger for large-scale discoveries. I also consider the effect of information disclosure, and find that, for repeat sales observations, property sales that occur following a discovery within 200 meters *and* after public registers are available observe a sizable loss in value. The robustness each of these findings, however, appears questionable, since their magnitude and statistical significance are sensitive to model specification.

Chapter 2. Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data[2]

2.1 Introduction

In energy economics and policy, it is important to understand how the demand for an energy input—such as electricity or natural gas—changes when the price of that input changes. This information, which often conveniently summarized into a price elasticity, namely the percentage change in demand when the price changes by 1%, allows regulators to estimate the welfare effects experienced by consumers as the regulatory environment is changed, as utilities enter or exit a market, and adequately plan infrastructure and grid investments (Labandeira et al., 2012).

The price elasticity of demand is also a key determinant of the tax revenue, effectiveness, and the burden falling on the shoulders of electricity generators, industry and consumers in the presence of a carbon tax, the remedy that has been put in place in some countries (e.g. Australia, the UK, and Sweden[3]) to encourage a shift away from fossil fuel usage and the associated CO_2 emissions (e.g., Hammar and Sjöström, 2011; Mori, 2012).

Given much recent interest and policy focus on improving energy efficiency, some observers have voiced concern over the rebound effect, namely the increase in energy use due to the fact that improved energy efficiency lowers the price per unit of energy services. The rebound effect erodes the efficiency gains and, if sufficiently pronounced, may even offset them completely. The key parameter for predicting the end outcome of improvements in energy efficiency is the elasticity of energy demand with

[2] This chapter was co-authored with Anna Alberini.
[3] For full list of countries, see http://www.carbontax.org/ (last accessed 31 August 2015). In the case of the UK, a climate levy imposes a tax on fossil fuels, but not as a direct function of its carbon content.

respect to efficiency, and this in turn can be shown to be equal to the negative of the price elasticity of demand, minus one (Sorrell and Dimitropoulos, 2008; Sorrell, 2007). Based on this identity and available estimates of the price elasticity on demand, Sorrell and Dimitroupolos conclude that the rebound effect in residential energy use is relatively small.[4, 5]

Price elasticity estimates for residential electricity demand vary widely across the economic literature. Alberini et al. (2011) review a number of studies, and suggest that differences might be due to the sample period, the type of data used (panels, pseudo-panels, cross-sections, time series), geography, and level of aggregation of the data. In more recent studies, the price elasticity of electricity consumption ranges from as low as -0.06 (Blazquez et al., 2013) to as high as -1.25 (Krishnamurthy and Kriström, 2013). In general, it is assumed that the price elasticity of demand for electricity is low; a meta-analysis by Espey and Espey (2004), for example, reports that the median short run elasticity for 36 studies is -0.28. From the output of the National Energy Modeling Systems model, the Energy Information Agency estimates short-run elasticities of -0.12 to -0.21 and a long-run elasticity of -0.40 when projecting residential energy demand over 25 years under different electricity and natural gas price scenarios (U.S. Energy Information Administration, 2014). These low elasticities imply limited fuel switching and result in relatively small changes in the numbers of electric furnaces, air-source heat pumps, and gas heating equipment.[6]

---

[4] By contrast, Davis (2008) uses actual energy use measurements in a randomized controlled trial featuring high-efficiency clothes washer to show that the rebound effect is negligible.

[5] Gillingham et al. (2013) deploy a similar approach and arrive at similar conclusions with cars and driving.

[6] Price elasticities of commercial and industrial demand are in relatively scarce supply. Mori (2012) surveys the literature and reports a handful of studies, with long run price elasticities ranging between -

The purpose of this paper is to systematically investigate the possible causes of such large variation. This is not mere intellectual curiosity, given the importance of the price elasticity of demand in utilities infrastructure planning, and energy and environmental policy analysis.

We examine seven possible factors that may explain why there is so much variation in the estimate of the price elasticity of electricity demand. The first possible reason is the period over which the elasticity was estimated, which in earlier research has spanned from one year (Krishnamurthy and Kriström, 2013) to over 40 years (Dergiades and Tsoulfidis, 2008). Another is whether over that period the price of electricity was rising or falling. Previous studies have examined this hypothesis, although they are limited to macro data (Gately and Huntington, 2002; Ryan et al., 1996). A third issue concerns the level of data aggregation, a recurring subject of concern as a source of bias (Bohi, 1981; Blundell et al., 1993; Blundell and Stoker, 2005). Aggregation reduces the variation in price, a key factor in identifying its elasticity of demand, and conceals the heterogeneity across more disaggregated units.

When the data used for estimating residential energy demand are a panel, another important issue is the degree to which unobserved heterogeneity is accounted for, along with the associated matter of variation in price. Does most of the variation in price come from within units over time, or is it primarily occurring between the units? We expect the

---

0.32 and -1.37 for commercial demand, and -0.22 and -0.83 for industrial demand. Lim et al. (2014) use 41 years' worth of data from Korea and produce estimates of -0.42 (short run) and -1.01 (long run) for the service sector, For Japan, Hosoe and Akiyama (2009) estimate commercial and industrial elasticities between -0.09 and -0.30 (short run) and -0.12 and -0.56 (long run), with more elastic demand in rural areas. Kamerschen and Porter (2004) report commercial and industrial elasticities of -0.34 to -0.55.

“within” estimator typically used with fixed-effects models to perform poorly in the presence of low variation within units over time.

Comparing estimates between studies is further complicated by how prices are measured. Studies in this field are typically subject to endogeneity of price and/or measurement error. First, prices are not always available at the individual household level (Alberini et al., 2011). Second, many studies are forced to rely on average price paid per kilowatt-hour (kWh) even though the original pricing structure faced by the household is a two-part tariff or block pricing. This makes price endogenous with consumption.[7] Endogeneity and/or measurement error can be addressed using instrumental variable (IV) estimation, and the success of this procedure depends crucially on the availability and quality of the instruments.

Finally, we consider the detail of the information available about the household or the dwelling. In recent years, several papers (Auffhammer et al., 2014; Ito, 2014; Allcott, 2011; and Allcott and Rogers, 2014) have deployed panel datasets provided by utilities, with electricity usage readings at a high level of granularity, but virtually no information about the household or the home, despite the importance of behavioral aspects and of the structural characteristics of the dwelling in influencing consumption patterns.[8]

To help reconcile the differences in price elasticity estimates that pervade this literature, we examine if, and how, each of these seven issues may be playing a role. We

[7] The matter is even more complicated in the presence of different pricing schemes for each household, as is the case when special offers, discounts etc. are introduced. Langer and Miller (2013) discuss the importance of manufacturing pricing and discounts in the case of car sales, showing that model estimation results and the coefficient(s) on price change dramatically when such discounts and special offers are controlled for.

[8] Alberini and Towe (2015) show that the effects of replacing certain types of electricity-using equipment are captured more sharply when one conditions on past usage *and* home characteristics.

use three public, nationwide datasets from the U.S. – the American Housing Survey (AHS), the Residential Energy Consumption Survey (RECS), and the Energy Information Agency's (EIA) Forms EIA-861. The former two provide information about electricity consumption at the household level, and the latter about total sales (in Kwh) for each class of customers, including residential consumers.

To see how sensitive price elasticities are to the sample period, we exploit the panel nature of the AHS and EIA datasets, create cross sections for each wave, and run regressions for each year. We also construct a fixed effects model for the full panel dataset, and include in our regressions the price interacted with a dummy denoting whether the price has been rising or otherwise relative to the previous year. The price elasticities from these regressions are shown to be relatively stable over time. Our results also suggest that whether the prices are increasing or otherwise makes little difference on the estimates of the price elasticities.

Next, we explore the issue of aggregation bias by first estimating models using the micro data, and then by aggregating electricity usage, prices, etc. to the metropolitan area or state level for the AHS, and state level for Forms EIA-861. The evidence from our experiment is mixed. With the AHS, more aggregation results in a more inelastic demand, whether or not we include fixed effects for the cross-sectional unit being considered in that run. This finding is similar to that of Halvorsen and Larsen's (2013), who show that the effect of aggregation depends crucially on the distribution of price and household income in the sample. With the Forms EIA-861 data, more aggregation produces a more elastic demand, although this effect is smaller when our models include cross-sectional-unit fixed effects.

We also examine how the price elasticities change when we control for unobserved heterogeneity. With both the AHS and Forms EIA-861 datasets, including cross-sectional-unit fixed effects always produces a more inelastic demand function. One reason for this result might be the limited "within" variation, as the "between" variation in price across our units always trumps the "within unit" variation.[9] Another reason might be that omitting the fixed effects biases the price elasticities.

To examine the effect of instrumental variables we work with the micro-level data for both the AHS and Forms EIA-861 datasets, and our instrument is the state-level price of electricity. However, the price elasticities from Two Stage Least Squares (2SLS) are quite similar to their Ordinary Least Squares (OLS) counterparts. Should this be caused by instruments of poor quality, we turn to the 2009 RECS, where we experiment with two alternate instruments. The first is the average state-level price, and the second is a cross-validation (CV) instrument, where the instrument for each household's price is the average of the price faced by all other households in the same state. The results are similar across instruments, and indicate a plausible and large change when going from OLS to 2SLS.

Finally, from the 2009 RECS, we find that including housing characteristics and capital stock in the model makes the demand more price inelastic. This is consistent with our expectations, since controlling for capital stock produces short run price elasticities.

[9] This is typical with data from the US, where the presence of different jurisdictions (states and counties) almost always ensures greater variation across jurisdictions than over time within a jurisdiction. Datasets from other countries may exhibit the opposite features.

The remainder of this paper is organized as follows. Section 2.2 reviews the literature. Section 2.3 discusses our empirical methodology, and Section 2.4 describes our datasets in more detail. We present our results in Section 2.5. Section 2.6 concludes.

2.2 Previous literature

Theory and empirical work distinguish between short-run and long-run price elasticities. Short-run elasticities hold the stock of energy-using durables fixed. Long-run elasticities come from models where the stock of capital is allowed to adjust. These models include i) static models that omit capital stock, ii) dynamic panel models motivated by a partial adjustment model (Houthakker, 1980), and iii) combinations of discrete choice and continuous-variable models that explain how households choose fuels or equipment and how much electricity (or other energy inputs) they use simultaneously (Nesbakken, 2001). Examples of dynamic models and the resulting long-run elasticities include Bernard et al. (2011); Paul et al. (2009); Dergiades and Tsoulfides (2008); Maddala et al. (1997); Alberini and Filippini (2011); Bernstein and Griffin (2006); Alberini et al. (2011); Okajima and Okajima (2013); and Boogen et al. (2014). In these models, the long-run estimate is equal to the short-run elasticity (i.e., the coefficient on log price) divided by one minus the coefficient on the lagged dependent variable.

Alberini et al. (2011) summarize estimates of the price elasticity of demand from earlier studies. We provide an updated assessment of select studies in Table 2.1. Clearly, the variation in the price elasticity is substantial. The absolute value of estimates range from near zero to over one; a majority fall below 0.5. Possible reasons for such variation may include each study's geography, study period, and hence variation in prices.

To illustrate, some studies are more limited in terms of geography, such as Garcia-Cerutti (2000), which considers county-level data only for the state of California. Others are limited in the number of time periods. Reiss and White (2005), for example, use household-level, cross- section data for California for 1993 and 1997, and Boogen et al. (2014) use Swiss household-level, cross sectional data for 2005 and 2011. The estimates of price elasticity may also be influenced by the type of data that are used. Most studies in Table 1 use panel data to exploit the variation in price within units over time. However, several remain confined to cross sections or time series (Kamerschen and Porter, 2004; Dergiades and Tsoulfides, 2008).

Based on Table 2.1, it would seem that in recent years sample sizes have grown as a result of longer time series and wider cross sections, and data have been more available at more disaggregated levels such as the state (Paul et al., 2009; Alberini and Filippini, 2010), census block group (Borenstein, 2009), or household (Boogen et al., 2014; Fell et al., 2014; Alberini et al., 2011). In the absence of true panels, Bernard et al. (2011) construct a pseudo-panel using four waves of survey data from Quebec that results in only 108 observations.

Table 2.1. Selected empirical studies and price elasticity estimates

Study	Type of data coverage	Variables included								Estimate(s) for fuel demand
		Cl	I	P	lag	HH	S	C	O	
Quigley and Rubinfeld (1989)	Annual Housing Survey, cross section, 1980	x					x	x	x	−0.1 energy
Maddala et al. (1997)	State-level, panel data, 1970–1990	x	x	x	x					−0.19 to − 0.21 *short run* (− 0.56 to − 1.03 *long run*) electricity
Metcalf and Hasset (1999)	RECS household-level, rotating panel, 1984, 1987, and 1990	x	x	x		x	x	x		−0.78 to − 1.11 electricity
Garcia-Cerrutti (2000)	California county-level, panel data, 1983–1997	x	x	x						Long-run: −0.17 electricity
Kamerschen and Porter (2004)	Nationwide total, time series, 1973–1998	x	x	x					x	Long-run: −0.94 to − 0.85 electricity
Bernstein and Griffin (2005)	State-level, panel data, 1997–2004	x	x	x	x				x	−0.243 (− 0.32) electricity
Reiss and White (2005)	California RECS, household-level, multi-year cross sections, 1993 and 1997	x	x			x	x	x	x	−0.85 to − 1.02 electricity
Dergiades and Tsoulfidis (2008)	Nationwide total, time series, 1965–2006	x	x	x	x				x	−0.386 short-run (− 1.06 long-run) electricity
Borenstein (2009)	California Census block group-level, panel data, 2000-2006	x	x						x	-0.12 to 0 (marginal price); -0.95 to 0 (average price); -0.46 to 0.09 (expected marginal price)
Paul et al. (2009)	State-level, panel data, 1990–2006	x	x	x	x				x	−0.13 (− 0.36) electricity
Alberini and Filippini (2010)	State-level, panel data, 1995-2007	x	x	x	x	x				−0.15 to − 0.08 (− 0.78 to − 0.44) electricity

Ito (2014)	California household level, panel data, 1999-2008								x	−0.09 (marginal price); -0.12 to -0.11 (average price); -0.1 (expected marginal price)
Alberini et al. (2011)	AHS household-level, panel data, 1997,2007	x	x	x	x	x	x	x		−0.736 (−0.814) electricity
Fell et al. (2014)	CEX household-level, panel data, 2006-2008	x	x			x	x	x		−0.82 to − 1.02 electricity
Studies outside the U.S.										
Nesbakken (1999)	Norway household level, multi-year cross-sections, 1990, 1993,1994, 1995	x	x			x	x	x		-0.24 to -0.53 energy
Bernard et al. (2011)	Quebec household-level, cross-section, 1989–2002	x	x	x	x					−0.51 (− 1.32) electricity
Gans et al. (2013)	Ireland household level, cross section, 1991-2009	x	x			x	x	x		-0.93 to -0.44 electricity
Blazquez et al. (2013)	Spain province-level, panel data, 2000-2008	x	x		x	x			x	-0.07 (-0.19) electricity
Okajima and Okajima (2013)	Japan prefecture level, panel data, 1990-2007	x	x		x					−0.397 (−0.487) electricity
Krishnamurthy and Kriström (2013)	OECD household level, cross section, 2011		x			x	x	x		−1.25 to -1.18 (based on samples that use all of the countries covered by the study) electricity
Blazquez et al. (2013)	Spain province-level, panel data, 2001-2009	x	x						x	−0.09 to -0.06 electricity
Boogen et al. (2014)	Switzerland household level, cross section, 2005, 2011		x			x	x	x	x	-0.59 to -0.54 (-0.65 to -0.68) electricity

Notes: for variables included, the following acronyms are used: C, climate; I, income; P, price of other substitutes; Lag, lag terms for price and/or quantity; HH, household level information; S, housing structure information; C, capital stock; O, other information (e.g. time fixed effects)

The question of the importance of rising or falling prices during the study period has received some attention in the literature. Haas and Schipper (1998), Gately and Huntington, (2002), and Ryan et al. (1996) examine this issue at the national level with time series data, and support the notion that investments in energy efficiency during periods of higher prices may result in little sensitivity to price changes when prices decline.

The issue of aggregation and aggregation bias is well studied in the econometrics literature (Blundell et al., 1993; Blundell and Stoker, 2005; and Stoker, 1993). From Table 1, we note that with a few exceptions (e.g., Kamerschen and Porter, 2004), studies with highly aggregated data appear to produce more inelastic demand functions than their disaggregated counterparts. However, the effect of aggregating is not necessarily monotonic. Bohi (1981) explains that bias is likely to occur with aggregation, but that the direction and magnitude of the bias is unknown a priori.

In the context of electricity, aggregation has been most directly examined by Halvorsen and Larsen (2013), who find that price elasticities estimated from aggregated micro data and macro data are lower than those estimated directly from micro data, and in some cases even positive. Using the framework of a quasi Almost Ideal Demand System model ("quasi" because consumption, rather than budget share, is used as the outcome variable), they demonstrate that the differences between the price elasticities from models that use data at different degrees of aggregation are attributable to the distribution of prices and income in the sample. Aggregating relies on a strong set of assumptions about homogeneity in terms of consumer preferences and price exposure: one cannot simply aggregate consumers who observe different prices and respond differently to price changes, and expect to find stable price elasticities.

Estimates of the price elasticity of demand may also vary across studies because of the specific set of covariates included in the model. Specifications that control for energy-using capital are generally interpreted as providing short-run elasticities. With panel data that lack information about capital stock, housing and equipment in the home, the assumption is made that these are approximately constant over time and are captured into household-, home- or meter-level fixed effects (Ito, 2014; Auffhammer, 2014). We interpret price elasticities from these types of studies to be short-run elasticities.

2.3 Methods

In this paper we focus on static models and short-run elasticities. Our research questions are summarized in Table 2.2. The ideal approach to answer most of them is to use one or more micro-level panel dataset(s) to estimate residential energy demand functions. Cross sections extracted from such panels can be used to examine issue 1 in Table 2.2. We exploit the panels to test the effect of rising or falling prices on elasticity (item 2 in Table 2.2). We increase the level of geographic aggregation of the observation (e.g., to the city or state level) to get total demand or the demand for a representative consumer to study the extent of aggregation bias (item 3 in Table 2.2).

Comparisons between fixed-effects models and pooled-data specifications allow us to cast light on issues 4 (unobserved heterogeneity) and 5 (within variation in electricity prices and the associated identification issues) of Table 2.2. Instrumental variable estimation is used to explore the extent of endogeneity bias (issue 6), whereas more detailed specifications that control for household and dwelling characteristics and equipment are used to explore issue 7.

Table 2.2. Summary of research questions

Issue	Data to be used	Methodology
1. Are price elasticity estimates influenced by the sample period?	AHS, Forms EIA-861	Produce and plot a series of cross sectional estimates for each year covered by each dataset
2. Are price elasticity estimates influenced by whether prices are rising or falling?	AHS (state level), Forms EIA-861 (state level)	Create a second variable which interacts price with an indicator for whether the price at the state level is higher or lower than in the previous year
3. Does aggregation bias price elasticity estimates?	AHS	Compare estimates obtained using the same data aggregated at the levels of household, city, and state
4. What is the effect of controlling for unobserved heterogeneity in panel datasets?	AHS, Forms EIA-861	Compare pooled-data and fixed-effects specifications
5. Variation in price: within and between variation	AHS, Forms EIA-861	Produce and compare summary statistics on prices
6. Endogeneity of price (due to measurement error, use of average or marginal price in the presence of block pricing or two-part tariffs, etc.)	RECS	Compare OLS with IV estimation
7. How sensitive are the elasticities to the inclusion of variables controlling for capital stock?	AHS, RECS	Compare specifications with and without such controls

Throughout this paper, we fit a log-log demand function, where the dependent variable is *lnQ*, the log kilowatt-hours (kWh) of electricity consumed in year *t*. The independent variables are lnP^e, the log price per kWh; *lnI*, log household income; *lnHDD* and *lnCDD*, log heating and cooling degree days (HDDs and CDDs), respectively; and lnP^g, log price of natural gas per cubic foot.[10] Where appropriate, we also add terms ***HH***, ***D***, and ***C***, namely vectors of household level characteristics, dwelling characteristics, and capital stock. This is a static demand function because it does not include lagged consumption (or price) in the right-hand side of the equation.

When we use the full panel of data, the regression equation is

$$(1)\ lnQ_{it} = \beta_0 + lnP^e_{it}\beta_1 + lnI_{it}\beta_2 + lnHDD_{it}\beta_3 + lnCDD_{it}\beta_4 + lnP^g_{it}\beta_5 + \boldsymbol{HH_{it}}\beta_6 + \boldsymbol{D_{it}}\beta_7 + \boldsymbol{C_{it}}\beta_8 + \gamma_t + \delta_i + \varepsilon_{it},$$

and this includes time fixed effects γ_t and sectional-unit fixed effects δ_i, where *i* denotes the unit. We define the cross-sectional unit *i* in different ways in different runs. When we use the AHS (RECS) data in their original form, *i* denotes the dwelling (household) and so regression equation (1) describes log dwelling (household) consumption in year t. When we use the Forms EIA-861 in their original form, *i* is a utility-state pair and the dependent variable in equation (1) is log sales to residential customers. With Forms EIA-861, equation (1) is amended to:

$$(2)\ lnQ_{it} = \beta_0 + lnP^e_{it}\beta_1 + lnHDD_{st}\beta_2 + lnCDD_{st}\beta_3 + lnP^g_{st}\beta_4 + \delta_i + \gamma_t + \varepsilon_{it},$$

since dwelling and household characteristics are irrelevant. Note that *HDD*, *CDD*, *and* P^g are available only at the level of the state *s*, but not at the utility service territory level.

[10] Natural gas is an obvious substitute to electricity for heating purposes.

With the AHS, we experiment with aggregating the original micro-level data to the metro area or the state level. We do so in two ways. Specifically, we compute either i) metro-area or state-level total consumption, and then take the log of it, or ii) compute the average usage per household in the metro area or at the state level, and then take its log transformation. When we follow procedure i), we add log N_i in the right-hand side of equations (1), where N is the number of households in the geographical area over which the total is calculated. This is because, all else the same, geographical areas with more numerous households will result in larger total consumption. With either procedure, the cross-sectional unit i denotes the metro area or state.[11] With Forms EIA-861, we also compute either i) the state-level total consumption, and then take the log of it, or ii) compute the average usage per customer in each state, which we then log.

2.4 Data

Candidate datasets for our analyses are micro-level, panel datasets that document energy usage and price, and ideally user characteristics, with sufficient geographical coverage and over a sufficiently long study period. We have identified two such datasets—the AHS and the Forms EIA-861. Information about these two datasets is summarized in Table 2.3, and the AHS is described in detail in Alberini et al. (2011).

[11] Technically, in the case where we create the average consumption per household in a specified geographical area, one should run weighted least squares because averaging the original consumption figures over uneven groups of consumers creates groupwise heteroscedasticity. However, when one obtains data already in aggregate form from government agencies, the size of the groups is most often unknown, and for that reason weights cannot be deployed. In this paper we assume that this is indeed the case and run unweighted regressions.

The AHS collects information from the household living at a specified residence every two years. The unit of observation is the home, not the household, and so when a household moves out of a home and is replaced by another, the AHS is administered to the latter. With this panel dataset, one can use fixed effects at the household or house level. House-specific fixed effects imply slightly longer panels than with household specific fixed effects.

In the AHS, information is collected about the dwelling structure, tenure and ownership, and associated costs (e.g., rent or mortgage payments), plus household sociodemographics, energy bills, heating and cooling and stock of energy-using appliances. Detailed geographical identifiers are provided only if the household resides in a metropolitan area with 100,000 or more people, and so in this paper attention is focused on 55 metro areas across the US. We use the data from 1997 to 2009. As in Alberini et al. (2011), prices are imputed at the city level to derive annual electricity consumption in kWh. Prices were obtained from state public utility commissions, utilities' web pages, etc., and represent Standard of Service tariffs, not necessarily the marginal price schedule by a given household.

Table 2.3. Description of datasets used

	Residential Energy Consumption Survey (RECS)	American Housing Survey (AHS)	Energy Information Agency (EIA), Forms EIA- 861
Who collects the data	EIA	HUD	EIA
Data type	Repeated cross sections	Rotating panel	Panel
Geographic identifiers	State* or state group	Metropolitan Area (MSA)	State
Frequency of data collection	Every 4 years	Every two years (with some MSAs additionally surveyed in between)	Annual
Universe	US households	US homes	US electrical utilities
Number of cross-sectional units used in this paper	6,070	16,947 housing units, 23,011 households	5,844
Household characteristics available?	Yes	Yes	No
Structural characteristics of the home available?	Yes	Yes	No
Inventory and vintage of capital stock?	Yes	Major capital stock only; no vintage	No
Energy consumption data	Total annual household electricity consumption (kWh), and total cost paid ($)	Total annual household cost ($). Electricity consumption derived by dividing through price.	Total annual sales (kWh) and revenue ($)

Endogeneity concerns for energy price variable	Measurement error due to use of average price; also, block pricing makes consumption correlated with price	Measurement error due to the use of average price, which is measured using the price of the utility provider associated with household's MSA	Measurement error due to use of average price; also, block pricing makes consumption correlated with price

*The sixteen states individually identified in RECS 2009 are Massachusetts, New York, New Jersey, Pennsylvania, Illinois, Michigan, Wisconsin, Missouri, Virginia, Georgia, Florida, Tennessee, Texas, Colorado, Arizona, and California; for years prior to then, only California, Florida, New York, and Texas.

Forms EIA-861 data are a panel. The cross-sectional unit is a utility in a given state. If a utility serves more than one state, we regard each utility-state pair as a separate cross-sectional unit. For each such unit, we have sales and revenues for each year by class of customers. For consistency with our AHS data, we use data from 1997 to 2009. This gives us 5,210 unique utilities, and 5,844 unique cross-sectional units (as a utility may serve more than one state). Prices are computed as revenue divided by kWh sold.

We also have a cross-sectional dataset—RECS—which provides information about a household's composition and stock of appliances, and reports electricity consumption at the annual level. We use the 2009 wave only, which contains 12,083 observations. The state of residence is exactly identified 16 states (listed in table 2.3). For some observations, we only know that the household is located in one of a group of two or more states.[12]

The nationwide coverage in the 2009 RECS is close to that of our AHS sample: 77.6% of the observations are located in one of the 14 states uniquely identified in RECS.

[12] For example, one such state group includes Alaska, Hawaii, Oregon, and Washington. RECS does not identify the city or county of residence of the household.

Observations for two of the remaining 16 RECS states (Virginia and Missouri) do not appear in our AHS sample. Prices are computed as electricity bills divided by consumption. They are thus the average price per kWh paid by a household.

The datasets used in this paper are limited to the 48 contiguous states (no Alaska or Hawaii[13]), and apply a number of data cleaning criteria.[14] A close comparison of the three datasets (see Table A.1 in Appendix 1) suggests good comparability in terms of the variables of primary interest to us. In the year 2009, we find the average consumption per household to range between 11,497 and 12,804 kWh, and the average price of electricity to be 10 to 12 cents per kWh (in 2009 dollars).[15]

Our household level datasets (RECS and AHS) exhibit similar overall variation in consumption and prices, and are likewise similar in terms of heating and cooling degree days. Other statistics, however, suggest interesting differences between the datasets. Single family homes in RECS, for example, are on average 36% larger in terms of square footage. While the availability of air conditioning is similar across the two datasets, the fraction of homes which primarily rely on electricity for heating is higher in RECS (29%) than in AHS (24%). The geographic distribution of the samples is also different. In terms of Census Divisions, our AHS sample is more heavily concentrated in the Mid-

[13] In the case of RECS, Alaska and Hawaii are not individually identified. For this dataset, however, we thus excluded observations which A) were identified with the state group comprising Alaska, Hawaii, Oregon, and Washington, and 2) observed a value of zero for heating degree days or cooling degree days.

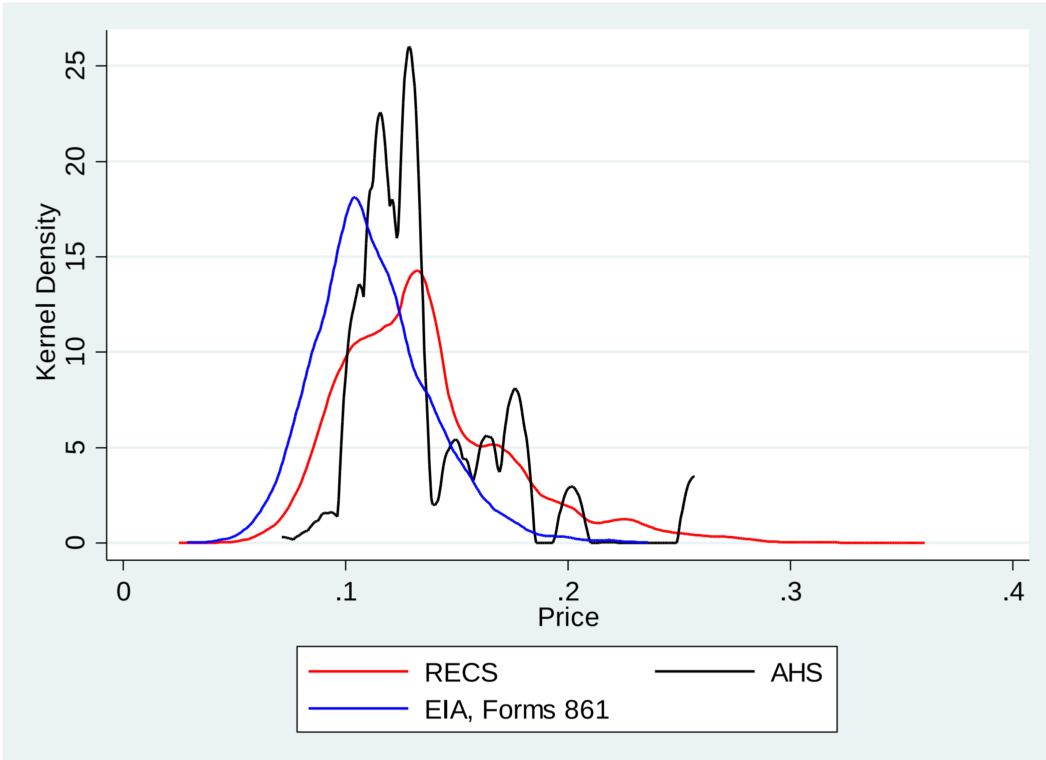
[14] For the AHS, we impose the same restrictions from Alberini et al. (2011). For RECS, which has 12,056 observations, we exclude homes 1) which are not single family homes (attached or detached) (3,379 observations), 2) with square footage less than 400 or over 10,000 square feet (32); 3) with heating or air conditioning that is applied to multiple units (79); 4) with an average price of less than one cent or over 50 cents per kWh (4); 5) with a renewable energy on-site system (163); 6) with imputed values for square footage, income , number of household members, the year the home was built, the type of heating and cooling equipment used, and indicators for an attic and basement (2,202); or 7) have a heating or cooling degree day value of zero (5).

[15] These statistics are based on samples that exclude the top and bottom 1% of the distribution of kWhs used. All regressions are based on such trimmed samples.

Atlantic, Midwest, West South Central, and Pacific divisions, whereas our RECS sample has larger shares in the Northeast, West North Central, South Atlantic, and East South Central divisions (see Figure A.1 in Appendix 1). These differences are due to the combination of different sampling frames across the two surveys, plus the fact that we retain for our purposes only the AHS observations from metropolitan areas with over 100,000 people, as all others lack sufficiently detailed geographical identifiers.

In terms of prices, in Figure 2.1 we plot the kernel density distribution of prices for observations in the year 2009 for each of our three datasets. We observe the smoothest distribution among the average prices from Forms EIA-861, followed by RECS, which has a somewhat longer upper tail. The price distribution for the AHS, in contrast, is quite "jagged" by comparison. These different distributions may explain the different price elasticities of demand when the data are aggregated to a coarser geographical level, as shown by Halvorsen and Larsen (2013) with data from Norway.

Figure 2.1. Distribution of prices from each dataset, for year 2009



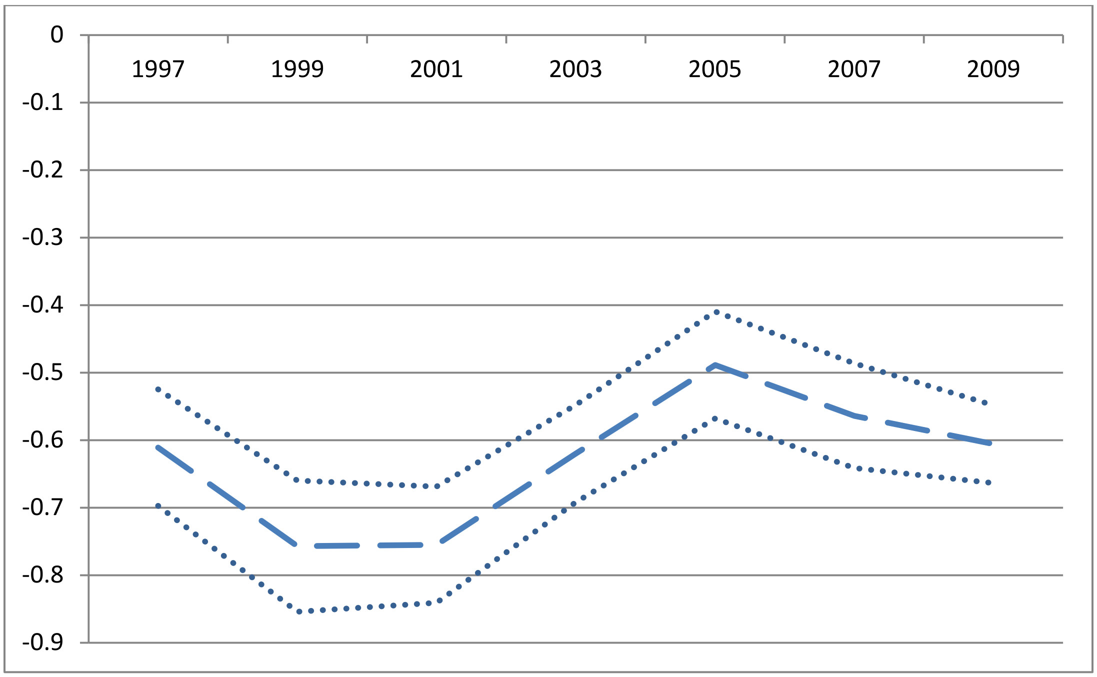
2.5 Results

2.5.1 The effect of timing and sample periods

We begin with item 1 from Table 2.2, by examining the sensitivity of price elasticity estimates to the sample period. To do so, we exploit the panel nature of the AHS and EIA datasets and create cross sections for each year, which we use to run separate, year-by-year regressions. The regressions for Forms EIA-861 are based on creating a representative residential customer for each utility-state. We include Census Division dummies in the AHS regression, and state-level dummies for Forms EIA-861.

Price elasticities from the regressions for the AHS, and their 95% confidence intervals, are plotted in Figure 2.2, and range from -0.76 to -0.56. This variation from some years to others is consistent with that observed by Nesbakken (1999) for Norway in 1990, 1993, 1994, and 1995. Her energy price elasticities are -0.24, -0.57, -0.33 and -0.53, respectively.[16]

Figure 2.2. Price elasticities: the effect of using repeated cross sections to obtain yearly estimates (source: AHS, household level)



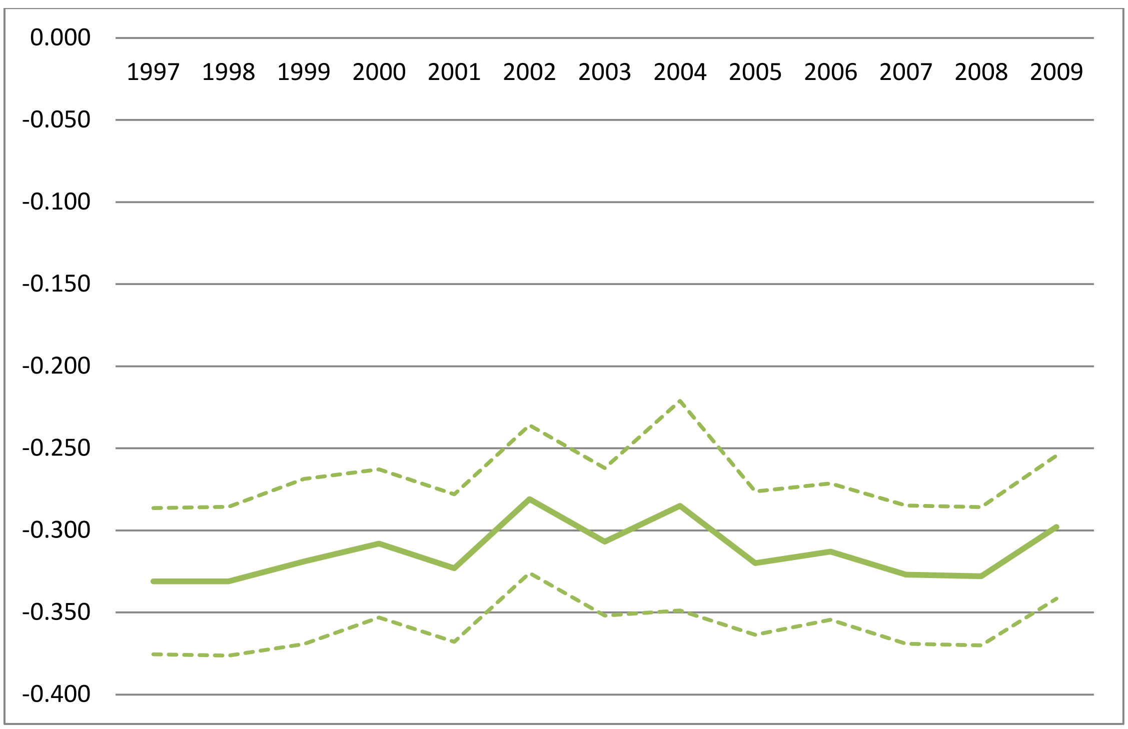
Note: estimates are provided with 95% confidence interval. Additional regressors include: the log price of gas; log household income; log cooling degree days; log heating degree days; log square footage; house age; house age squared; number of rooms; number of floors; log number of household members; indicators for home ownership, electric heating, gas heating, fuel oil, central AC unit, and window AC units; and Census Division dummies.

We perform the same exercise for the Forms EIA-861 data. The price elasticities are plotted in Figure 2.3. These price elasticities are much lower than those calculated

[16] Nesbakken (1999) uses a discrete-choice/continuous model where households choose their fuel combinations (e.g., electricity only, electricity plus heating oil, electricity plus gas, etc.) and total energy use a function of the fuel prices, housing characteristics and household sociodemographics. The model takes the equipment as given, but allows for substitution between fuels. Also see Nesbakken (2001).

for the AHS, in Figure 2.2. This may be due to the different extent of measurement error in prices in this dataset, and/or to the fact that we have included state dummies in the Forms EIA-861 regressions, which forces the estimation to rely on variation in price within a state. The price elasticities in Figure 2.3 are relatively stable over time, although they appear to be estimated imprecisely due to the smaller sample sizes (max. 3,036 observations per year), as indicated by the wide confidence bands at the 95% confidence level.

Figure 2.3. Price elasticities: the effect of using repeated cross sections to obtain yearly estimates (Source: EIA, Form 861, utility level)



Note: estimates are provided with 95% confidence interval. Additional regressors include: the log cooling degree days (state level), log heating degree days (state level), and state dummies.

Do these yearly fluctuations depend systematically on how high the prices are, or on how much price variation across units we observe every year? To answer this question, we run three regressions for each set of estimated price elasticities. In the first regression, we regress estimated price elasticity on average price in that year. In the second, we regress the estimated price elasticities on the variance of price across units, and in the third on the coefficient of variation of price (standard deviation across units divided by average price).

The results of these regressions are reported in Table 2.4. The results are mixed. The AHS panel of Table 2.4 suggests an inverse relationship between elasticity and average price in that year, and a positive relationship between price elasticity and variance of price, but neither relationship is statistically significant. The coefficient of variation is positively related to the price elasticity. With Forms EIA-861, the relationship between price elasticity and price variance, and that between price elasticity and the coefficient of variation of price, is negative. Average price remains negatively associated with the price elasticities.

Table 2.4. Price elasticities: the effect of average price and variance on annual estimates

	American Housing Survey			Forms EIA-861		
Average price	-4.286			-1.331		
	(8.406)			(1.247)		
Price variance		310.8			-96.79	
		(162.3)			(56.93)	
Coefficient of variation			2.672*			-1.241**
			(1.253)			(0.557)
Constant	-0.107	-1.128***	-1.508**	-0.175	-0.249***	-0.005
	(1.025)	(0.263)	(0.413)	(0.129)	(0.0379)	(0.138)
Observations	7	7	7	13	13	13
R-squared	0.049	0.423	0.477	0.094	0.208	0.311

Notes: standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

It is difficult to draw conclusions from this exercise. It is possible that households might be more sensitive to price changes if or where the prices are especially high, implying a causal relationship between price levels and responsiveness to price. It is also possible that price elasticities are more reliably estimated, and are truly small, when the variation in price is greater. Ultimately, however, it is difficult to say if price elasticities have systematically changed over time, given the fluctuations in the point estimates observed from year to year, the width of the confidence intervals around them, and the limited number of data points from which to draw inference.

For item 2 of Table 2.2, we test for asymmetry of price elasticities between periods of rising and declining prices by fitting fixed effects models on the original, full panel datasets and adding the log price interacted with a dummy variable denoting whether the price has risen or declined relative to the previous period.

Our approach differs from previous literature, which used country-level time-series data. Ryan et al. (1996) model price asymmetry by regressing fuel expenditure shares on the current period price, $ln(P_{j,t})$ (*j* referring here to commodity *j*), which is approximated by the decomposed value $ln(P_{j,t-1}) + (\Delta P^{+}/P_{j,t-1}) + (\Delta P^{-}/P_{j,t-1})$,[17] where ΔP_j^{+} and ΔP_j^{-} respectively denote increases and decreases in prices. Alternatively, the approach used by Haas and Schipper (1998) and Gately and Huntington (2002) decomposes price for each time period *t* into four separate terms: 1) the log price during the very first period *t*=1, 2) the cumulative increases in log maximum price up to period *t*,

[17] The full formula, which this term approximates for small changes in *P*, is $ln(P_{j,t}) = ln(P_{j,t-1}) + ln(1 + (\Delta P_j/P_{j,t-1}))$.

3) the cumulative increases in sub-maximum log price up to period *t*, and 4) the cumulative decreases in log price up to period *t*.

We opt for the dummy variable approach for two reasons. First, unlike Haas and Schipper (1998) and Gately and Huntington (2002), we do not have within-year price variation. Second, the length of our panels is too short. With our data, the number of lagged prices per unit (house or utility) ranges from zero to six for the AHS, and zero to 12 for Forms EIA-861.

The results from this exercise are summarized in Table 2.5. To save space, in Table 2.5 we report only the coefficients for the two main price variables of interest—log price and log price interacted with the increasing-price dummy—which are reported in the first and second columns, respectively. The first four rows report the results for the AHS at different levels of aggregation: the first two are at the household level, with the first controlling for housing unit fixed effects and the second controlling for household fixed effects; the third and fourth are for average household consumption at the metropolitan area and state levels, respectively. Rows five and six report the results for representative consumer models using Forms EIA-861 at utility and state levels, respectively.

Column (2) of Table 2.5 shows that, in all cases, the coefficient on the interaction term of interest is very small. Only in two of the cases—with the AHS housing unit fixed effects model, and the Forms EIA-861 utility-level models—is the coefficient statistically significant at the conventional levels, and, even so, it remains practically insignificant.

Table 2.5. Price elasticities: the effect of controlling for rising/declining prices

		(1)	(2)	(3)
Dataset	Cross sectional unit	Log price (per kwh)	Log price x Dummy variable (equals 1 if average price exceeds previous period's)	Proportion of sample with dummy variable equal to 1
AHS	Housing unit	-0.609*** (0.0505)	0.00667* (0.00372)	0.47
AHS	Household	-0.671*** (0.0541)	0.00359 (0.00393)	0.47
AHS	City	-0.390*** (0.0905)	0.00655 (0.00755)	0.5
AHS	State	-0.356*** (0.0995)	0.00723 (0.00919)	0.7
Forms EIA-861	Utility	-0.214*** (0.0124)	-0.00237*** (0.000436)	0.43
Forms EIA-861	State	-0.275*** (0.0171)	-0.00189 (0.00135)	0.74

Notes: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. All models are aggregated at, and include fixed effects for, the level specified in parentheses. For the AHS, the previous period price is the average price from the previous time surveyed, two years prior; for the EIA, the previous period is one year prior. For the household- and housing unit-level regressions with the AHS, additional regressors include: the log price of gas; log household income; log cooling degree days; log heating degree days; log square footage; house age; house age squared; number of rooms; number of floors; log number of household members; indicators for home ownership, electric heating, gas heating, fuel oil, central AC unit, and window AC units, and year dummies. For the city-level (state-level) regressions with the AHS, additional regressors include: the log average price of gas at the city-level (state level), log average heating and cooling degree days at the city level (state-level), and year dummies. For the utility- and state-level regressions with Forms EIA-861, additional regressors include: the log price of gas (state level), log HDDs and CDDs (state level), and year dummies.

Insufficient variation in “rising prices” does not appear to be a contributing factor to these results, as the third column in Table 2.5 suggests that rising prices account for between 47 percent and 74 percent of the observations. For both the AHS and Forms EIA-861 dataset, this proportion increases with the observations’ level of aggregation.

2.5.2 Aggregation

The results from our analysis of aggregation bias issues (item 3 of Table 2.2) are displayed in Table 2.6a and 2.6b. In Table 2.6a, rows 1) and 2) report the price elasticities from using the AHS panel in its original form. The two rows differ in that row 1) is based on a model with dwelling-specific fixed effects, while row 2) is based on a model with household-specific fixed effects. We get the coefficients in rows 4) and 5) from models where the dependent variables are average metro area consumption, and an average state-level consumption, respectively. The model in column (A) omits unit-specific fixed effects; that in column (B) includes them.

With Forms EIA-861, we first fit models to the utility-level sample (row (3)) and then at the state-level sample (row (5)). In each case, we consider both pooled data and fixed effects models (columns (C) and (D), respectively).

Table 2.6a. Price elasticities: the effect of controlling for unobserved heterogeneity - Representative consumer model (Dependent variable: log kWh per consumer)

		(A)	(B)		(C)	(D)	
		American Housing Survey			EIA Forms 861		
	Aggregated to…	Pooled data model	Fixed effects model		Pooled data model	Fixed effects model	
(1)	Household level	-0.611*** (0.0136)	-0.587*** (0.0300)	(dwelling fixed effects)			
(2)	Household level	-0.611*** (0.0136)	-0.619*** (0.0337)	(household fixed effects)			
(3)	Utility level				-0.308*** (0.00648)	-0.201*** (0.0115)	(utility fixed effects)
(4)	City level	-0.470*** (0.0509)	-0.436*** (0.0749)	(city fixed effects)			
(5)	State level	-0.470*** (0.0875)	-0.360*** (0.0881)	(state fixed effects)	-0.692*** (0.0337)	-0.285*** (0.0165)	(state fixed effects)

Notes: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. To allow for comparability between datasets, additional regressors for all regressions include: the log price of gas, log cooling degree days, log heating and cooling degree days, year dummies, and Census Division dummies. For the household-, housing unit-, and city-level specifications with the AHS, the price of gas, and heating and cooling days are measured at the city level. For the state-level regression with the AHS, the gas price and heating and cooling degree days are averaged at the state level. For both the utility- and state-level regressions with Forms EIA-861, the price of gas, and heating and cooling days are state averages. In these models, the dependent variable is the log average kWh consumed per consumer. Rows 1) and 2) report the price elasticities from using the AHS panel in its original form. Row 1) is based on a model with dwelling-specific fixed effects. Row 2) is based on a model with household-specific fixed effects. Rows 4) and 5) are based on models where the dependent variables are average metro area consumption, and an average state-level consumption, respectively. The model in column (A) omits unit-specific fixed effects; that in column (B) includes them. Columns (C) and (D)_are for the Forms EIA-861. Row (3) is for the utility-level representative consumer and row (5) is for the state state-level representative consumer.

Table 2.6b. Price elasticities: the effect of controlling for unobserved heterogeneity - Aggregated model (Dependent variable: log total kWh)

		(A)	(B)		(C)	(D)	
		American Housing Survey			EIA Forms 861		
	Aggregated to…	Pooled data model	Fixed effects model		Pooled data model	Fixed effects model	
(1)	Household level	-0.611*** (0.0136)	-0.587*** (0.0300)	(dwelling fixed effects)			
(2)	Household level	-0.611*** (0.0136)	-0.619*** (0.0337)	(household fixed effects)			
(3)	Utility level				-0.324*** (0.00650)	-0.203*** (0.0113)	(utility fixed effects)
(4)	City level	-0.466*** (0.0530)	-0.448*** (0.0753)	(city fixed effects)			
(5)	State level	-0.408*** (0.0900)	-0.354*** (0.0877)	(state fixed effects)	-0.740*** (0.0350)	-0.245*** (0.0181)	(state fixed effects)

Notes: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. To allow for comparability between datasets, additional regressors for all regressions include: the log price of gas, log cooling degree days, log heating and cooling degree days, year dummies, and Census Division dummies. For the household-, housing unit-, and city-level specifications with the AHS, the price of gas, and heating and cooling days are measured at the city level. For the state-level regression with the AHS, the gas price and heating and cooling degree days are averaged at the state level. For both the utility- and state-level regressions with Forms EIA-861, the price of gas, and heating and cooling days are state averages. In these models, the dependent variable is the log aggregated kWh consumed. Rows 1) and 2) report the price elasticities from using the AHS panel in its original form. Row 1) is based on a model with dwelling-specific fixed effects. Row 2) is based on a model with household-specific fixed effects. Rows 4) and 5) are based on models where the dependent variables are total metro area consumption, and total state-level consumption, respectively. The model in column (A) omits unit-specific fixed effects; that in column (B) includes them. Columns (C) and (D) are for the Forms EIA-861. Row (3) is for the utility-level total sales and row (5) is for the state state-level total sales.

For the AHS, using more highly aggregated observations produces a more inelastic demand. This applies whether or not we include fixed effects in the model. This pattern is consistent with that observed by Halvorsen and Larsen (2013), who even obtained positive elasticities based on data aggregated to the national level. With the Forms EIA-861 data, however, aggregating the utility-level data to the state level makes the demand more elastic. In the case of the fixed effects models, the difference between the estimates of these two models is substantial—almost 50%. That we observe different results from the AHS and EIA datasets seems plausible. As noted by Halvorsen and Larsen (2013), the relationship between the level of aggregation and the price elasticity is not necessarily monotonic, as the effect of aggregation depends to a large extent on the distribution of prices and income among the population of interest.

In Table 2.6b, we fit similar models as in Table 2.6a, but this time the dependent variable is the total amount of electricity consumed in the specified geographical area (rather than the average per consumer in that area).[18] For example, if the unit of observation is the state, the dependent variable is the log of total consumption in the state. The model specifications are identical to those in Table 2.6a, except that we control for the log number of consumers. In the case of the AHS, neither the city- nor state-level results change much compared to their counterparts in Table 2.6a. The Forms EIA-861 data displays similar results. In one case (the state fixed effects model), the coefficient is -0.245, which is within 20% of the coefficient (-0.203) from the same fixed effects model at the utility level (also shown in Table 2.6b).

[18] The first two rows of Tables 2.6a and 2.6b are thus the same.

2.5.3 Unobserved heterogeneity

Turning next to the effect of unobserved heterogeneity (item 4 of Table 2.2), we compare our pooled data models with their fixed effects counterparts. For simplicity, we comment on the results in Table 2.6a (the results in Table 2.6b are similar). For any given level of aggregation, consumers are more price inelastic when the model includes fixed effects. Aggregating the data increases the discrepancy between estimates from the pooled and fixed effects models. This effect is more pronounced with the Forms EIA-861 data.

With the AHS, the proportional change ranges from 15% (household-level model and household fixed effects) to 34% (state-level representative consumer model). With Forms EIA-861, the decrease ranges from 35%, for the utility-level representative consumer model, to 59%, for the state level representative consumer model. One reason for these findings may be that the variation in price is now limited to the “within” variation: as shown in Table 2.7, the “between” variation in price across units always trumps the former. Another reason might be that omitting the fixed effects biases the price elasticities.

Table 2.7. Summary statistics for price, compared within and between datasets

Dataset	Cross sectional unit	Years	Mean	S.D., overall	S.D., between	S.D., within
AHS	Housing unit	1997-2009, biennially	0.1218	0.0400	0.0382	0.0096
AHS	Household	1997-2009, biennially	0.1218	0.0400	0.0386	0.0084
AHS	City	1997-2009, biennially	0.1171	0.0378	0.0362	0.0118
AHS	State	1997-2009, biennially	0.0981	0.0300	0.0256	0.0163
Forms EIA-861	Utility	1997-2009	0.1035	0.0260	0.0251	0.0099
Forms EIA-861	State	1997-2009	0.0905	0.0254	0.0214	0.0139

2.5.4 The effect of instrumental variable estimation

In the context of our analysis, instrumental variable estimation (Item 6, Table 2) is justified for two reasons. First, it addresses the endogeneity of price where block or two-part tariffs are present. Second, it addresses measurement error in price. Measurement error is almost certainly present in our datasets. In general, it is difficult to find good-quality instruments, since we do not have marginal prices or policy changes that might be regarded as exogenous shocks and are correlated with price changes.

In what follows, we use two different instruments:[19] the first is the average price of electricity at the state level (provided by the Energy Information Agency); the second

[19] An alternative approach for constructing instruments, proposed by Mansur and Kahn (2013), would be to use the price of inputs used to generate electricity, expressed as a linear combination of each input commodity's share of the total, multiplied by each commodity's price. Mansur and Kahn use the utilities' capacity data for three commodities: oil, natural gas, and coal. This offers the advantage of having an instrument that is excluded from the electricity demand equation, unlike the weather; it is limited, however, by its omission of price and share data on nuclear fuel, hydroelectric power, and other fuels. We experiment with a similar approach, using data on the resource mix of oil, coal, natural gas, and nuclear energy at the state-level from Forms EIA-906, and data on the prices on each commodity from the EIA (for coal, Forms EIA-423 and EIA-923.; for oil, the WTI spot price; for natural gas, the Henry Hub spot price; for uranium, the weighted average price by owners and operators of US civilian nuclear power reactors from the Uranium Marketing Annual Report). While this alternative approach covers more input, it too could be problematic to the extent that net generation shares are also a function of weather. Our approach has

is a cross-validation (CV) instrument, namely the average of one's neighbor's price. We reason that a given household's neighbors face a similar tariff structure as that household, but the prices they face cannot possibly cause this household's consumption. We use the state price as an instrument with both the AHS and RECS, and the CV approach only with RECS. The "neighbors" are all other households in the same state.[20] Because it is essential to know which state a household lives in, we are forced to use only the RECS observations for which the state is exactly identified (3,919 observations out of the over 6,000 available to us after applying our data cleaning criteria).

We begin with Panel A in Table 2.8, which displays the results for fixed effects models, using the AHS and Forms EIA-861 at levels of aggregation below the state. The first three rows refer to the AHS data, which are used with two different types of fixed effects (house and household fixed effects, respectively) and after aggregating them to the city level.

In rows 1-2, the IV estimation procedure produces a lower price elasticity than OLS, but within at most 24% of the price elasticity from OLS or the within estimator. The results are qualitatively similar when we consider the results from a representative consumer model aggregated at the city level, as shown in the third row of Panel A. In this case, going from OLS to 2SLS changes the price elasticity estimate by 27%.

---

mixed results. For some runs, like those at the survey-level for the AHS, the instrument performs comparably to those using state-level prices as instruments (for those with housing unit fixed effects, the elasticity is -0.526 [standard error: 0.157]; with household fixed effects, -0.466 [0.160]). At the other extreme, for some results like RECS 2009, it performs poorly (-2.284 [1.722], when capital stock is controlled for; -2.675 [1.648] otherwise).

[20] We do not attempt to apply instrumental variable estimation with the Forms EIA-861 dataset. This is because the most obvious instrument, state-wide price, is computed from the Forms EIA-861 themselves and would thus overlap with the other variables in the regression.

Table 2.8. Price elasticities: the effect of using instrumental variables (IV)

	OLS	IV
Panel A. Panel data, using fixed effect models at the level specified in parentheses		
AHS (survey level with housing unit fixed effects)	-0.587***	-0.449***
	(0.0300)	(0.0403)
AHS (survey level with household fixed effects)	-0.619***	-0.493***
	(0.0337)	(0.0444)
AHS (aggregated to city level)	-0.456***	-0.333***
	(0.0784)	(0.110)
Panel B. Cross sectional data		
RECS – controlling for capital stock and equipment (2009,	-0.354***	-0.638***
IV=avg state price)	(0.0308)	(0.0463)
RECS – controlling for capital stock and equipment (2009,	-0.354***	-0.604***
IV=CV estimate)	(0.0308)	(0.0448)
RECS – no capital stock and equipment controls (2009,	-0.435***	-0.735***
IV=avg state price)	(0.0323)	(0.0478)
RECS – no capital stock and equipment controls (2009,	-0.435***	-0.698***
IV=CV estimate)	(0.0323)	(0.0461)

Notes: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. For all regressions - except for RECS "CV" estimates - the instrumental variable used consists of the average price calculated at the state level, using the data from Form 861. For RECS, results are confined to 16 states, as residents from only those areas are identified at the state level. "CV" refers to "cross validation", whereby the instrument is constructed by taking the average price of all other households within the same state. For the household- and housing unit-level regressions with the AHS, additional regressors include: the log price of gas; log household income; log HDDs and log CDDs; log square footage; house age; house age squared; number of rooms; number of floors; log number of household members; indicators for home ownership, electric heating, gas heating, fuel oil, central AC unit, and window AC units, and year dummies. For the city-level (state-level) regression with the AHS, additional regressors include: the log average price of gas at the city-level (state level), log average HDDs and log average CDDs at the city level (state-level), and year dummies. For the regressions with RECS, additional regressors include: a dummy for the ownership of gas; an interaction term for gas ownership times the price of gas; log household income; log HDDs and log CDDs; log number of household members; indicators for home ownership; log square footage; indicator for home being attached unit; number of rooms; categorical dummies for year built; categorical dummies for stories; and indicators for a basement and an attic. Additional controls for RECS specifications controlling for capital stock and equipment include indicators for a central AC unit, one or more window AC units, and electric heating.

Should these findings be due to having poor quality instruments, we turn to the 2009 RECS, which allows us to experiment with both types of instruments. We display these results in Panel B of Table 2.8, where we consider two specifications—one with and one without controls for capital stock and equipment, which lead to short-run and long-run elasticities, respectively.

Overall, the results are similar across instruments, and indicate a large and plausible change in the price elasticity when going from OLS to 2SLS. This change is somewhat sensitive to the inclusion of capital stock in the model. For specifications with controls for capital stock and equipment, it ranges from 71 to 80 percent; for those without controls, 60 to 69 percent.

2.5.5 The effect of controlling for housing characteristics and capital stock

For our last question (item 7 of Table 2.2), we further examine the importance of controlling for housing characteristics and capital stock, when estimating price elasticity. Here, we focus exclusively on the RECS survey in 2009 – this time, using observations from all contiguous states (rather than the 16 states individually identified in the data). Table 2.9 summarizes the results from four models. Column (1) consists of a basic household demand model, by controlling only for the log price, log household income, state/state group dummies, and log heating and cooling degree days; columns (2), (3), and (4) further control for household characteristics (log household members, and an indicator for home ownership), characteristics of the housing unit (age of home, the log square footage, etc.), and capital stock (ownership of electric heating, a central AC unit, and a window AC), respectively. Table 2.9 illustrates that adding both housing characteristics and capital stock decreases price elasticity. This is consistent with the notion that controlling for capital stock produces price elasticities that should be interpreted as short run. The effect is modest, however.

Table 2.9. Price elasticities: the effect of controlling for household and dwelling characteristics, and appliance stock (Source: RECS, year 2009)

	(1)	(2)	(3)	(4)
log electricity price	-0.491***	-0.458***	-0.447***	-0.387***
	(0.0388)	(0.0354)	(0.0334)	(0.0322)
log household income	0.175***	0.118***	0.0660***	0.0594***
	(0.00896)	(0.00833)	(0.00785)	(0.00759)
Includes…				
state-group dummies	yes	yes	yes	yes
heating and cooling degree days	yes	yes	yes	yes
household characteristics	no	yes	yes	yes
housing characteristics	no	no	yes	yes
capital stock	no	no	no	yes

Notes: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Additional regressors for household characteristics (column (2)) include log number of household members and a home ownership dummy. Controls for housing characteristics (column (3)) include log square footage; an indicator for attached housing unit; number of rooms; categorical dummies for year built; categorical dummies for stories; and indicators for a basement and an attic. Controls for capital stock (column (4)) include indicators for a central AC unit, one or more window AC units, and electric heating.

Importantly, Table 2.9 shows that controlling for these variables has a strong effect on income elasticity, which falls to nearly one third of the initial estimate – from 0.18 to 0.06. Higher household incomes are correlated with other regressors—the size of one's house, the possession of capital stock appliances, etc.—that have a direct effect on consumption.

2.6 Conclusions

An important question in energy economics and policy is how elastic demand is with respect to energy prices. In the case of residential electricity demand, knowing how the demand changes with changes in price is important when planning for grid infrastructure, assessing the effects on consumer welfare of deregulation and other changes in public utilities' regulatory environment (Nakajima and Hamori, 2010), and determining the extent of any rebound effects associated with policies that promote or require energy efficiency improvements. The price elasticity of demand figures prominently in the context of current proposals for introducing a carbon tax, as it is often assumed that the effect of any such a measure would be small (Ryan et al., 2011; Di Cosmo and Hyland, 2013, Nakata and Lamont, 2001, and Mori, 2012).

There is a wide range of estimated price elasticities of residential electricity demand, and this range is attributed to a variety of factors, usually data and locale-related.[21] We have examined a number of such factors in this paper. We have worked with three micro-level datasets from the U.S. Two are at the household level (the AHS and RECS) and one at the utility level (Forms EIA-861). Two (the AHS and Forms EIA-861) are panel, whereas RECS is a cross section (repeated every four years). Consumption levels and prices can be aggregated (i.e., summed and averaged, respectively) over specified geographic areas, and it is possible to extract cross sections from the two longitudinal datasets.

[21] In developing countries, where theft and loss of electricity are common, it may be difficult to reliably estimate a price elasticity of demand (Jamil and Ahmad, 2013; Katiyar, 2005; Gaur and Gupta, 2016), and rising prices may worsen this problem.

We do not find strong evidence of price elasticities systematically changing over time; estimate appear to fluctuate within a range of 35 percentage points. However, splitting our panel datasets into annual cross sections, we do observe a negative relationship between price elasticities and the average prices for that year. Whether prices are rising or falling appears to have little effect on our price elasticity estimates. We also find that aggregating our data can produce higher *or* lower price elasticities, depending on the dataset used, and that controlling for cross sectional unit fixed effects (with panel data) tends to result in a more inelastic demand function. The latter finding may be attributable to limited within price variation, relative to the between price variation, and/or to bias resulting from the omission of fixed effects.

We experiment with instrumental variables, finding that addressing price endogeneity (which is due to two-part or block tariffs, or measurement error) has a small but noticeable effect for each of our datasets, and appears to be largest in the case of RECS. Finally, we find that controlling for housing characteristics and capital stock makes a difference. Omitting these variables overstates the price elasticity.

In general, the three datasets we used produce price elasticities of demand ranging from -0.2 to -0.8, confirming the notion that household electricity demand is not very elastic. Changing the estimation technique, aggregating the data or selecting specific years from the panel dataset can double or halve the price elasticity, which remains less than one. This is well within the range used in much policy work and in sensitivity analyses to such work. For example, Mori (2012) discusses the quality of existing estimates, uses the average from a meta-analysis of earlier studies, and generates values from a normal distribution assuming a central value of -0.43 and a standard deviation of

0.3 (where 95% of the values would be between -1.018 and 0.158). He concludes that the predictions of the effect of a carbon tax are qualitatively as these and other parameters are changed. However, it is entirely possible that in a different policy exercise, a relatively small change in price elasticity leads the analyst to completely different policy conclusions and recommendations.

The analyses in this paper are not without limitations. We are estimating price elasticities using annual consumption data, and annual price data, and for that reason we fail to capture month-to-month fluctuations. We use panel data, and account for unobserved heterogeneity by including unit-specific fixed effects, but doing so prevents us from taking advantage of the variation in price across units. In some cases, applying the same estimation technique (or data aggregation) to two different datasets results in revisions of the "base case" estimates that go in opposite directions. In others, we are unable to conclude whether changes in price elasticities from one year to the next are due to the movements of prices, the higher or lower variation in prices within a year, or something else. Finally, there is also question of external validity: we focus on residential consumers in the contiguous US states, and it is not clear to us whether we would reach the same conclusions with commercial or industrial entities.

Chapter 3. How much CO_2 do rural drivers emit, relative to their urban counterparts? Evidence from the UK

3.1 Introduction

Personal transportation is the source of various externalities, including traffic accidents, congestion, and emissions of conventional air pollutants and greenhouse gases. It is of general interest to the policymaking community to characterize the source of these externalities, in order to determine how to structure policies targeting their reduction, and who will be affected.

Recent research has taken an interest in profiling the carbon footprint of households of a population to better understand who would bear the most incidence from proposed policies such as a carbon tax (Baiocchi et al., 2010; Thumim and White, 2008; Büchs and Schnepf, 2013; Gough et al, 2011; Lenzen et al., 2006; Weber and Matthews, 2008). Considering total emissions of a household, including home energy emissions and indirect emissions,[22] this literature identifies different footprints as a function of income (Baiocchi et al., 2010; Gough et al., 2011; Druckman and Jackson, 2008), household composition (Büchs and Schnepf, 2013; Gough et al., 2011), and education (Baiocchi et al., 2010).

One relevant distinction to be made is between rural and urban households. Given their differences in lifestyles, housing quality, access to energy resources, and transportation needs and options, there is strong reason to expect their carbon footprint to be different (Glaeser and Kahn, 2010). In the case of personal transportation, it is typically assumed that rural households produce a higher carbon footprint than their

[22] Indirect emissions are those resulting not from the fuel used for the consumption of such goods as transportation or electricity, but the production process of consumed goods such as food and furniture.

urban counterparts, since they must generally travel longer distances to reach a destination and have fewer mass transit options.[23] But is this necessarily the case? For example, if rural households must drive more, they may drive more fuel efficient vehicles. Alternatively, their lifestyle may be structured in a way that involves making fewer or shorter trips (e.g., telecommuting, commuting outside of the central city) (Glaeser and Kahn, 2004; Gillingham et al., 2015).[24]

In this chapter, I focus on the geographical distribution of carbon dioxide emissions from driving cars. Specifically, I focus on rural and urban households in the UK, using repeated cross sections from the UK National Travel Survey (NTS). I contrast driving behavior between these two groups, and ask three related questions: First, do rural households purchase more vehicles, and/or are their vehicles higher CO_2 emitters? Second, are their vehicles driven more? Third, how do their carbon emissions compare to urban households? These answers are relevant, for policymaking purposes, in identifying the incidence and impact of carbon-mitigating legislation on these respective segments of the population.

To answer these research questions, I first model the number of vehicles owned by a household, using a discrete choice framework. Using a similar framework, I then model the choice of CO_2 emissions category for new vehicle purchases, based on the amount of emissions generated by the vehicle. Conditioning on this selection in the style of Dubin and McFadden (1984), I then model the travel demand for newly purchased

---

[23] In the case of the US, the American Public Transportation Association (2012) and US Federal Highway Administration (2012) note that public transportation accessibility for rural communities is generally lacking. Other areas where public transportation in rural areas is recognized as a problem include Australia (Nutley, 2003) and Scotland (Velaga et al., 2012).

[24] The ambiguity concerning whether urban and rural regions generate more air pollution is discussed in greater detail by Nechyba and Walsh (2004).

vehicles, controlling for the vehicle's rural and urban status. Finally, I use these estimations to compute the contribution of rural and urban residents to total CO2 emissions.

Among vehicle-owning households, I find that rural households have a greater probability of owning any given number of vehicles greater than one. Their vehicles are also higher emitting, on average, although the difference is rather small. Rural households also drive 12.9 percent more miles with their vehicles than comparable urban households. Finally, I estimate an annual contribution of 3.4 tons of CO_2 from driving per rural household, which is 58 percent higher than an urban household. Collectively, rural emissions account for 27.7 percent of total CO2 emissions from personal vehicular transportation.

The remainder of this chapter is organized as follows. Section 3.2 reviews the literature. Section 3.3 discusses the empirical model, and Section 3.4 discusses our datasets in more detail. The results are presented in Section 3.5, followed by a discussion on the carbon footprints in Section 3.6. Section 3.7 concludes.

3.2 Previous literature

3.2.1 Emissions profiles of UK Households

Previous literature has extensively considered how carbon emissions vary among different types of households. Typically, this is done by comparing households on the basis of socioeconomic, demographic, and geographic differences in such outcomes of interest as home energy consumption, transport emissions, and indirect emissions. The UK is an often considered geography of study. Several works (Thumim and White, 2008; Büchs and Schnepf, 2013; Gough et al, 2011) use expenditure survey data over several years, which they supplement with other data to convert into CO_2 emissions. Thumim and White (2008), in simulating the distributional impact of a personal carbon trading system, observe higher emissions from rural households due to older and less efficient housing and heating systems, lower ambient temperatures from a lower heat island effect, and a higher use of oil, rather than gas. Transportation emissions, while higher in rural areas, account for a lower share of total emissions than urban households. Büchs and Schnepf (2013), in their assessment, find that home energy, transportation, and indirect emissions are all three positively associated with income, and that among low income, elderly, and workless populations home energy emissions produce the largest share of the three. Among rural households, they find higher emissions for all three types, including transport emissions which are 16 percent higher.[25]

Profile studies of this type are often limited in the amount of sociodemographic data, limiting the extent to which emissions can be attributed to such characteristics. Druckman and Jackson (2008), for example, divide household into seven “Supergroups,” seven geographic areas which the UK National Output Area Classification (OAC) defines on the basis of socio-demographic characteristics. While they identify higher energy consumption from “countryside” households, particularly from fuels other than those used for electricity and gas, they cannot determine which characteristics used by the OAC to define the groups (e.g. detached housing, 2+ cars per household, low population density) to attribute this finding to. Baiocchi et al. (2010) consider the heterogeneity of emissions profiles among UK households, by combining emissions data from the UK

[25] The source of these higher transport emissions is direct – i.e. from motor fuel, rather than public or flight transportation.

Environmental Accounts, which aggregates emissions for 76 geographic sectors in the UK, with neighborhood-level data on lifestyle characteristics from a residential classification database, Acorn. They observe higher emissions among higher income and lower educated populations. While they are unable to directly differentiate between rural and non-rural households, they do control for lifestyle variables associated with the rural regions, including "villages with wealthy commuters," "flourishing families," and "wealthy executives," which are associated with higher total emissions. They also find that emissions are positively correlated with other variables associated with rural households, including average household size and the average size of homes.

Concerning the question of carbon taxation, there is concern about regressivity. Although emissions tend to be higher among higher income households, expenditures as a share of income may still be higher for poorer households (Starkey, 2012; Grainger and Kolstad, 2010). Using grants and loans to subsidize improvements in energy efficiency has been discussed as one means of addressing this (Dresner and Ekins, 2006).

3.2.2 Emissions studies from other countries

Carbon footprint studies for households extend to other countries, as well. In the case of the US, Weber and Matthews (2008) combine information on industries and trading patterns with expenditure survey data to examine the role of increased trade and income on the distribution of emissions across households. Glaeser and Kahn (2010) take a different approach by examining the emissions profiles of specific cities across the US, combining household-level microdata with regional power plant data to create an index with which to compare cities. Within metropolitan areas, they report generally lower emissions among households in central cities, compared to peripheral suburbs. Across metropolitan regions, higher emissions are reported in areas with higher growth, and lower levels of land use controls. A similar effort to compare the emissions profiles of cities across China by Zheng et al. (2010) finds higher total emissions, and transportation emissions exclusively, among cities with colder climates and higher levels of income.

A multi-country approach is considered by Lenzen et al. (2006), using expenditure and input-output data to examine energy consumption from five countries (India, Denmark, Australia, Brazil, and Japan). Their measures for urbanity, which differ for each country, suggest that higher density regions produce lower emissions, with the exception of India where, they suggest, energy needs are higher due to the increased need for mobility and poor public transportation options.

3.3 The model

This chapter examines three research questions. First, do rural households tend to own more vehicles, and/or vehicles that are higher emitters? Second, do they drive these vehicles more miles than urban households? Third, how do the total vehicle emissions of rural households compare to those from urban households?

I begin with the first question, by modeling the choice for the number of vehicles owned by a household. To do so, I employ a discrete choice model, where the indirect utility for the number of cars k owned by a household i in time period t is a function of household characteristics, $\boldsymbol{H}$, including rural/urban status:

$$(3.1)\ V_{ikt} = \propto_k \boldsymbol{H_{it}} + \epsilon_{ikt},$$

where ϵ is an i.i.d. extreme value error term with scale 1. Households, when determining the number of vehicles to own, choose the number which offers the greatest utility. The probability of choosing each number is then

$$(3.2)\ \Pr(k) = \exp(\propto_k \boldsymbol{H_{it}}) / [\textstyle\sum_w^W \exp(\propto_w \boldsymbol{H_{it}})].$$

This is a multinomial logit model, with α_1 normalized to zero for identification.

To study the emissions of rural and urban households, I also consider vehicle level purchases, and specifically the household's choice of a vehicle's emissions class. For each new purchase, a household's indirect utility can be expressed as a function, once again, of household characteristics, $\boldsymbol{H}$:

$$(3.3)\ V_{ijt} = \propto_j \boldsymbol{H_{it}} + \varepsilon_{ijt},$$

where ε is an i.i.d. extreme value error term with scale 1. Households, when making a purchase, choose the vehicle type which offers the greatest utility, with the probability of choosing such being

$$(3.4)\ \Pr(j) = \exp\left(\propto_j \boldsymbol{H_{it}}\right) / [\textstyle\sum_m^M \exp(\propto_m \boldsymbol{H_{it}})].$$

As with equation (3.2), this too is a multinomial logit model.

For my second research question, I consider a demand equation for travel of the form

$$(3.5)\ lnVMT_{ivjt} = \beta_0 + \beta_1 lnP_{vjt} + \beta_2 I_{it} + \beta_3 \boldsymbol{H_{it}} + \beta_4 \boldsymbol{x_{vjt}} + \theta_t + \eta_{ivjt}$$

where *lnVMT* denotes the log vehicle miles traveled (VMT) by household *i*, using vehicle *v* of type *j* in time period *t*, *lnP* the log price per mile, and *I* the household income. The price per mile is calculated by multiplying the price (pence) per liter of fuel by the vehicle's fuel consumption rate, measured in terms of liters per 100 km. Here, $\boldsymbol{x}$ is a vector of vehicle characteristics of *v*. Year of travel is controlled for using fixed effects θ. The UK National Travel Survey (NTS) does not disclose the specific make and model of each vehicle, so I cannot include make and model effects.

Unobserved factors contributing to a household's selection of *lnVMT*, captured by η, are likely to also be correlated with their choice of *lnP*, part of which - since *P* includes fuel economy - is determined at the time of vehicle purchase (Gillingham, 2011; West, 2004; Goldberg, 1998). For example, households which anticipate a high VMT may be more likely to purchase a vehicle with better fuel economy, or larger, more comfortable vehicles. It is necessary, therefore, to account for this relationship by separately modelling the selection process associated with *lnP*.

One approach might be to instrument for price per mile, using a two-stage least squares (2SLS) framework. Alternatively, one can choose to model the choice of the vehicle. Ideally, we would like to model the choice of the specific vehicle *v* (Gillingham, 2011; Kiso, 2013). The NTS's lack of information on the vehicle make and model, however, renders this approach infeasible in my case. I instead opt to model the choice of vehicle type *j*, capturing a household's broad preference for a vehicle's size class.

The two-step approach employed here is that suggested by Dubin and McFadden (1984), known as the "Conditional Expectation Correction Method." To do this, I allow for possible correlation between η and ε, making the choice of vehicle's emissions category and the subsequent VMT endogenous. To obtain consistent estimates of the coefficients in the VMT equation, I condition the outcome on the choice of the emissions category, from equation (4.4). In Dubin and McFadden, it is assumed that

(3.6) $E(\eta|\varepsilon) = \frac{\sigma\sqrt{6}}{\pi}\sum_{k=1}^{K} r_k(\varepsilon_k - E(\varepsilon_k))$

with r being the correlation coefficient between η and ε_k. They further illustrate that the coefficients in the VMT equation can be consistently estimated by running OLS on

$$(3.7)\ lnVMT_{ivjt} = \beta_0 + \beta_{1j} lnP_{vjt} + \beta_2 I_{it} + \beta_3 \boldsymbol{H_{it}} + \beta_4 \boldsymbol{x_{vjt}} + \theta_t + \sum_{k \neq j} \propto_k \left(\frac{\hat{P}_{ikt} * \ln(\hat{P}_{ikt})}{1-\hat{P}_{ikt}}\right) + \propto_j \hat{P}_{ijt} + \eta_{ivjt}$$

where k denotes vehicle types not chosen by i, the $\hat{P}$'s denote predicted probabilities of selecting each vehicle type from the choice model outlined above for emissions type, and the α's are the correlation coefficients from equation (3.6), rescaled by the standard deviation. Note that, in equation (3.7), I allow for the price elasticity of demand for mileage to vary with each choice.

Equation (3.7) differs slightly from Dubin and McFadden's original model, which imposes the constraint that the α coefficients sum to zero. According to Bourguignon et al (2007), who compare the performance of these two models, along with selection correction procedures by Lee (1983) and Dahl (2002), the modified version of Dubin and McFadden's model that relaxes this constraint is the most robust.

In selection models, it is a common practice to include variables in the selection stage which are not present in the second stage, so that estimation of α does not rely solely on functional form. In this chapter, in the multinomial logit equation I include the price of gas at the time of purchase (rather than the time of travel), and year fixed effects for the estimated year of purchase (rather than the year of travel).

3.4 Data

The models in this chapter rely on the UK NTS, which collects information on thousands of UK households each year. Households provide information on their basic sociodemographic characteristics, their means of transportation, vehicle ownership, and commuting behavior. For one week, they are also asked to maintain a travel diary. They also provide data for each of their vehicles, including the vehicle's size class, age, registration year, engine capacity, and the number of miles traveled in the preceding 12 months. As the NTS is a cross section repeated on an annual basis, I use data for the years 2002 to 2010.

Vehicle emissions are reported for only a subset of the data, about 34,000 vehicles. The main reason for this is the Vehicle Excise Duty (VED), which was originally enacted in 1999. In its original form, all vehicles were taxed the same amount; within a few months, however, the tax was differentiated for vehicles on the basis of engine capacity. It was not until March 1st, 2001 that the VED began to discriminate on the basis of measured emissions rates. For vehicles registered before this date, the VED was not retroactively enforced, and for this reason data on emissions are missing. As one of the main goals of this chapter is to compare the vehicle emissions between urban and rural households, this limits my analysis to cars registered on or after March 1, 2001. This has the result of making the sample size small in the first year of my study, with more and more vehicles added to it in the subsequent years. To make the sample composition of vehicles in each year of the survey more comparable to one another, and level the contribution of each survey year to the assessment, I limit the sample for the travel demand model to new vehicles, which I define as those aged two or less.

The price per mile for each vehicle is estimated using two pieces of information: the fuel economy, and the price of fuel. Neither is provided directly in the NTS dataset. To calculate the former, using a vehicle fleet database from the UK Vehicle Certification Agency, I estimate regressions for fuel economy, separately for diesel and gasoline, solely as a function of emissions, for different years. These coefficients allow me to estimate, with extremely high precision,[26] the fuel economy for vehicles in the NTS for which emissions data are provided.

Fuel prices are available at the level of the government office region (GOR), for the reason that the NTS does report the geographic information of households for lower levels of aggregation. The location, and size, of these GORs, are illustrated in Figure 3.1. Fuel prices are provided by the Automobile Association (AA) on a monthly basis for both unleaded-95 and diesel.[27] I calculate the fuel price to be matched to each household as an average of the fuel prices observed for the household's GOR during the preceding twelve months (including the month surveyed). Fuel prices are further distinguished on the basis of a vehicle's fuel type, such that the price averaged for diesel cars corresponds to diesel, rather than unleaded-95.

[26] In all cases, these regressions produced R^2 values exceeding 0.99. This is because emissions are proportional to the fuel consumption rate (the inverse of fuel economy).

[27] For fuel prices prior to February, 2005 nine of the eleven GORs were consolidated into three larger regions; in such cases, households are matched to the prices of these larger regions.

Figure 3.1 Government Office regions in Great Britain

Sources: for government office regions of England, University of Edinburgh ShareGeo Open spatial data repository; for Wales and Scotland, the UK Ordinance Survey.

In this chapter, the sample is restricted to vehicles classified as small cars, small/medium cars, medium cars, large cars, and “Land Rover, Jeep, or similar” cars (i.e., sport utility vehicles), thus omitting cars classified as light vans, “other,” or “NA” (210 vehicles, which is less than two percent of the sample). I also trim the remaining sample to eliminate outliers, dropping vehicles in the top and bottom one percent of the distribution of miles driven. I also eliminate cars with emissions exceeding 400 grams per kilometer (which translates to roughly 16 liters per 100 kilometers), or with an engine capacity exceeding 4000 cc – further reducing the sample by 36 observations. This leaves me with 10,469 observations.

A household’s geographic type is identified in the NTS from a categorical variable, delineating different types of urban areas. The UK Census defines rural settlements as those with populations below 10,000 (The Countryside Agency, 2004). To maintain consistency with the Census, rural households are defined in this paper as those residing in areas with a population of 10,000 or below.

Table 3.1 presents summary statistics for vehicle level observations in the sample, aged two or less. On average, they are driven 11,095 miles, which exceeds the average for older vehicles by over 2,500 miles. The standard deviation is also notably high, producing a coefficient of variation of 0.68. Most vehicles exhibit a fuel economy of less than 10 liters per 100 kilometers, but the distribution is right skewed. The fuel prices (which are 12-months averages) are more evenly distributed. The third row illustrates that, combining these two variables, drivers pay on average 5.87 pounds per 100 kilometers. Engine capacity, unsurprisingly, exhibits a similarly right skewed distribution as fuel economy. Vehicle age is almost evenly split between the ages of less than one versus one to two.

Table 3.1. Summary statistics for vehicle observations in travel demand model (n=10,469)

Variable	**Mean**	**S.D.**	**Min.**	**Max.**	**Mean among vehicle owning households (n=59.254)**
Miles traveled	11095.2	7588.14	700	45000	
Fuel economy (liters per 100 km)	6.70	1.43	3.33	15.97	
Fuel price (pence per liter)	87.96	8.21	76.23	110.11	
Price per 100 kilometers (pence)	586.58	123.53	328.09	1468.51	
Engine capacity (cc)	1703.90	444.04	599	3999	
Vehicle age					
Below one year	0.46	0.50	0	1	
1-2 years	0.54	0.50	0	1	
Size category					
Small car	0.14	0.34	0	1	
Small/medium car	0.32	0.47	0	1	
Medium car	0.30	0.46	0	1	
Large car	0.19	0.39	0	1	
Land rover/jeep	0.05	0.21	0	1	
Rural household	0.27	0.44	0	1	0.24
Household income quintile					
1	0.08	0.26	0	1	0.13
2	0.12	0.33	0	1	0.18
3	0.19	0.39	0	1	0.22
4	0.25	0.43	0	1	0.24
5	0.37	0.48	0	1	0.24
Household owns home	0.90	0.30	0	1	0.82
Household category					
1 or 2 adult household, with youngest child age 0-15	0.31	0.46	0	1	0.31
Family household, with adult child(ren)	0.15	0.36	0	1	0.13

Pensioner household	0.18	0.38	0	1	0.21
Other	0.37	0.48	0	1	0.35
Persons in household with driver's license					
1	0.25	0.43	0	1	0.40
2	0.64	0.48	0	1	0.52
3	0.09	0.28	0	1	0.06
4+	0.03	0.16	0	1	0.02
Time from bus to rail station >14 mins.	0.52	0.50	0	1	0.51
Walk time to rail station >26 mins	0.63	0.48	0	1	0.59
Walk time to bus stop >6 mins.	0.16	0.36	0	1	0.15
Socioeconomic group					
Professional/managerial	0.31	0.46	0	1	0.20
Clerical	0.25	0.43	0	1	0.24
Skilled manual	0.13	0.33	0	1	0.17
Other / other manual	0.06	0.24	0	1	0.09
Retired	0.23	0.42	0	1	0.26
Other economically inactive	0.02	0.13	0	1	0.03
Year					
2002	0.12	0.26	0	1	0.11
2003	0.14	0.27	0	1	0.11
2004	0.13	0.27	0	1	0.11
2005	0.11	0.25	0	1	0.12
2006	0.12	0.26	0	1	0.11
2007	0.11	0.25	0	1	0.11
2008	0.10	0.24	0	1	0.11
2009	0.08	0.22	0	1	0.11
2010	0.05	0.17	0	1	0.06
Government Office Region					
Northeast	0.05	0.22	0	1	0.04
NW & Merseyside	0.12	0.33	0	1	0.12
Yorkshire & Humberside	0.09	0.28	0	1	0.08
East Midlands	0.08	0.28	0	1	0.08
West Midlands	0.10	0.30	0	1	0.09
Eastern Region	0.10	0.30	0	1	0.10
Greater London	0.06	0.24	0	1	0.10
South East Region	0.15	0.35	0	1	0.15
South West Region	0.09	0.28	0	1	0.10

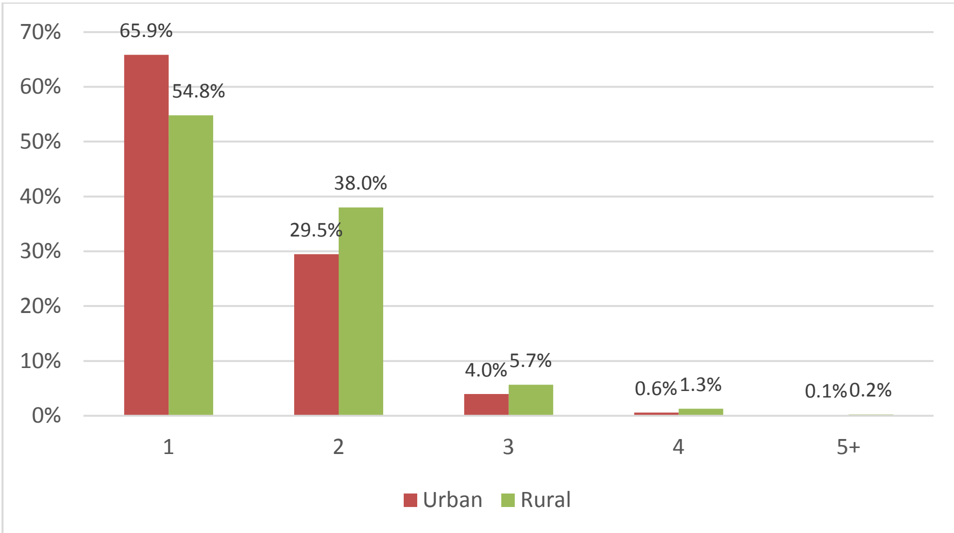
Wales	0.05	0.23	0	1	0.05
Scotland	0.10	0.30	0	1	0.09

Note: the term "vehicle owning households" excludes households with over five vehicles.

Table 3.1 also shows that most cars tend to be small/medium and medium in size, with smaller shares among the extremes, with larger vehicles and small cars. We also see larger shares of vehicles from households among the highest income quintiles, indicating higher vehicle ownership rates among such segments of the population. Higher shares of vehicle ownership are found among households which are homeowners, have families, are located further from public transportation, have a head member employed in a professional, managerial, or clerical profession, and were surveyed in the early 2000s. The geographic shares of the new vehicle sample appear to be generally similar to their shares for the sample for all households in the survey, with the most noteworthy exception being, unsurprisingly, Greater London, where ownership rates are much lower.

Vehicles from rural households comprise 27 percent of the sample. As rural households comprise 24 percent of the total sample of vehicle owners, as shown in the far right column of Table 3.1, this suggests a higher vehicle ownership rate among this group. Figure 3.2 takes a closer look at this, with a histogram for the number of vehicles owned by each type of household, with at least one car (thus excluding the 30.5% of urban households and 16% of rural households who do not own a car). Figure 3.2 indicates that rural households tend to own more vehicles. This finding is consistent with previous research from Thumim and White (2008) and Commins and Nolan (2010), for the UK and Ireland, respectively.

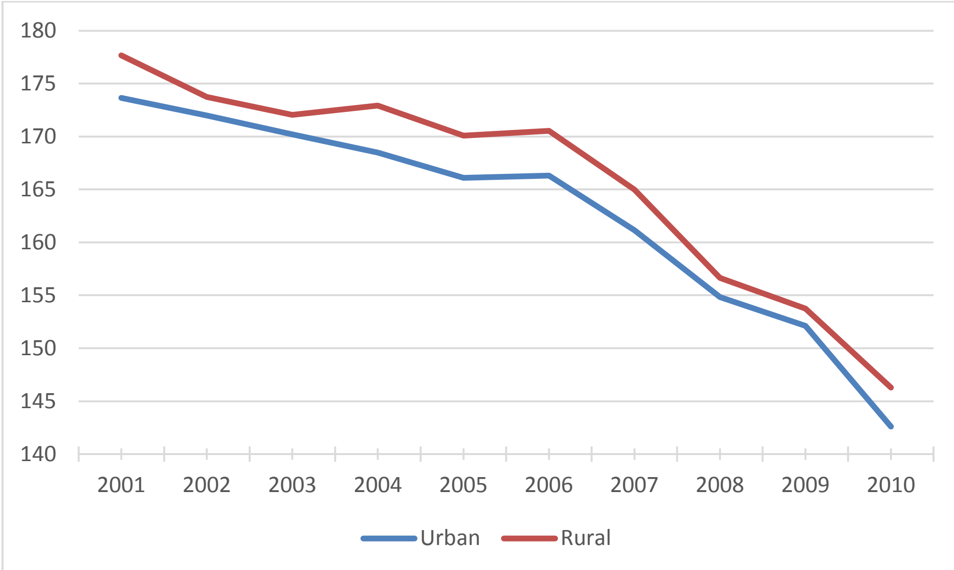
Figure 3.2. Number of vehicles owned by urban and rural households, among households with more than one vehicle



Source: UK National Travel Survey

In Figure 3.3, I consider the emissions profiles of new vehicles, by plotting the average emissions, per estimated year of vehicle purchase, for rural and urban households. Figure 3.3 tells us two things. First, from 2001 to 2010, for urban and rural households the general trend is negative; for the overall sample, emissions declined by 18%. This may be due to the fact that vehicles were becoming cleaner, and/or consumer preferences shifted towards cleaner cars. Second, rural households have higher-emitting vehicles. This difference is relatively small, though; emissions from rural cars are 1.5% higher emissions overall, and range from one to three percent in magnitude for any given year.

Figure 3.3. Average new vehicle emissions per year (grams per kilometer)



Source: UK National Travel Survey

3.5 Results

3.5.1 Number of vehicles owned

I start by considering the question of vehicle ownership, and how rural households differ in their demand for the number of vehicles. To answer this question, I consider all vehicles, rather than new cars only. While this includes a much larger share of the survey sample, including households that only have older cars, it still omits a sizable portion, 27.5%, of households which do not have a car. Among households with a vehicle, 99.9% own five or fewer vehicles. For this reason, I cap the choice set in my multinomial logit model at five vehicles, omitting from the sample any outliers exceeding that.

Table 3.2 provides the results for this multinomial logit model. In general, these results are fairly intuitive. Households with higher incomes who own their homes, are families, have children, have more drivers, are located further from bus stops and rail stations, and are of a managerial/professional socioeconomic group tend to own more than one vehicle. For many of these variables, such as income, the coefficients tend to be positive for choices corresponding to more vehicles – up to a point, at least. Large, significant coefficients for the last choice, 5 vehicles, are generally less common. As only 72, or 0.1%, of households own five vehicles, it is likely that identification for coefficients in this category proves more difficult.

Table 3.2. Multinomial logit results for number of vehicles owned (omitted category: 1 car)

VARIABLES	(1) 2 cars	(2) 3 cars	(3) 4 cars	(4) 5 cars
rural	0.353***	0.585***	1.145***	0.960***
	(0.0278)	(0.0581)	(0.122)	(0.272)
household income in 2nd quintile	0.0931*	0.0538	0.203	0.769
	(0.0490)	(0.129)	(0.326)	(0.831)
household income in 3rd quintile	0.300***	0.231*	0.342	0.929
	(0.0461)	(0.118)	(0.298)	(0.779)
household income in 4th quintile	0.619***	0.590***	0.956***	1.207
	(0.0464)	(0.117)	(0.289)	(0.776)
household income in 5th quintile	0.897***	1.260***	1.647***	1.726**
	(0.0485)	(0.119)	(0.295)	(0.787)
Household owns home	0.447***	0.707***	0.783***	1.649***
	(0.0347)	(0.0917)	(0.215)	(0.631)
Year 2002	-0.00164	0.402	0.862	0.371
	(0.102)	(0.248)	(0.614)	(1.044)
Year 2003	0.162**	0.584***	0.427	-0.351
	(0.0816)	(0.198)	(0.502)	(0.880)
Year 2004	0.224***	0.702***	1.392***	0.00326
	(0.0865)	(0.208)	(0.509)	(0.927)
Year 2005	0.271***	0.849***	1.359***	0.0216
	(0.0851)	(0.204)	(0.499)	(0.897)
Year 2006	0.236***	0.617***	1.284***	0.490
	(0.0853)	(0.205)	(0.496)	(0.856)
Year 2007	0.223***	0.770***	1.431***	0.161
	(0.0857)	(0.206)	(0.503)	(0.902)
Year 2008	0.244***	0.745***	1.172**	-0.00353
	(0.0860)	(0.205)	(0.502)	(0.907)
Year 2009	0.251***	0.952***	1.305***	0.321
	(0.0865)	(0.205)	(0.505)	(0.889)
Year 2010	0.412***	0.713***	1.185**	0.755
	(0.0954)	(0.226)	(0.544)	(0.929)
Year 2011	1.168	5.190	6.585	11.02
	(2.639)	(5.014)	(9.159)	(14.30)
Family household, with adult child(ren)	0.266***	0.439***	0.693***	0.653*

	(0.0409)	(0.0714)	(0.156)	(0.354)
Pensioner household	-0.639***	-0.806***	-1.264**	-14.64
	(0.0493)	(0.150)	(0.550)	(822.1)
Other household	-0.245***	-0.380***	-0.129	0.419
	(0.0278)	(0.0697)	(0.196)	(0.420)
2 persons in household with driver's license	3.238***	3.213***	2.232***	1.286**
	(0.0382)	(0.158)	(0.357)	(0.524)
3 persons in household with driver's license	4.095***	6.602***	5.768***	4.444***
	(0.0664)	(0.170)	(0.373)	(0.596)
4+ persons in household with driver's license	3.824***	7.269***	8.421***	7.257***
	(0.135)	(0.202)	(0.383)	(0.581)
Time from bus to rail station >14 mins.	-0.0188	-0.0408	0.139	-0.596**
	(0.0278)	(0.0599)	(0.133)	(0.298)
Walk time to rail station >26 mins	0.198***	0.269***	0.375**	0.296
	(0.0295)	(0.0642)	(0.146)	(0.309)
Walk time to bus stop >6 mins.	0.0856***	0.234***	0.329**	0.0702
	(0.0323)	(0.0669)	(0.138)	(0.332)
Socioeconomic group: clerical	-0.239***	-0.473***	-0.507***	-0.722**
	(0.0324)	(0.0652)	(0.143)	(0.332)
Socioeconomic group: skilled manual	-0.595***	-0.812***	-0.766***	-0.699**
	(0.0357)	(0.0724)	(0.151)	(0.324)
Other socioeconomic group / other manual	-0.563***	-0.948***	-1.003***	-1.101**
	(0.0459)	(0.101)	(0.220)	(0.510)
Socioeconomic group: retired	-0.638***	-1.212***	-1.099***	-1.015*
	(0.0476)	(0.107)	(0.235)	(0.518)
Socioeconomic group: other economically inactive	-0.514***	-1.021***	-1.083**	-0.399
	(0.0836)	(0.208)	(0.446)	(0.793)
NW & Merseyside	0.151**	0.305*	0.604	0.0816
	(0.0660)	(0.156)	(0.370)	(1.170)
Yorkshire & Humberside	-0.0610	0.173	0.240	1.022
	(0.0687)	(0.163)	(0.395)	(1.110)
East Midlands	0.0798	0.558***	0.856**	1.562
	(0.0695)	(0.159)	(0.374)	(1.084)
West Midlands	0.313***	0.780***	1.197***	2.020*
	(0.0682)	(0.157)	(0.366)	(1.057)

Eastern Region	0.156**	0.449***	0.948***	1.465
	(0.0666)	(0.156)	(0.360)	(1.063)
Greater London	-0.523***	-0.566***	0.0594	0.320
	(0.0693)	(0.166)	(0.380)	(1.104)
South East Region	0.308***	0.774***	1.138***	1.475
	(0.0642)	(0.149)	(0.353)	(1.059)
South West Region	0.130*	0.504***	0.879**	1.103
	(0.0669)	(0.157)	(0.369)	(1.110)
Wales	0.0539	0.484***	0.880**	1.359
	(0.0758)	(0.171)	(0.390)	(1.114)
Scotland	-0.109	-0.0661	0.000442	0.963
	(0.0689)	(0.164)	(0.390)	(1.084)
Constant	-3.930***	-7.805***	-11.09***	-11.94***
	(0.113)	(0.309)	(0.755)	(1.671)
Observations	59,254	59,254	59,254	59,254

Standard errors in parentheses; the symbols ***, **, * denote statistical significance at the levels of 1 percent, 5 percent, and 10 percent, respectively

Households surveyed in later years of the NTS tend to have more vehicles, as well. Households in Greater London – as would be expected - tend to own fewer cars, as do households in Scotland, the North East, Yorkshire, and Humberside. Households in rural regions are also predicted to own larger numbers of vehicles. The coefficients for all four vehicle choices are statistically significant for this variable, and positive. For the average observation in the sample, rural status translates to a 7% decrease in the probability of owning one car, and an increase in the probability of owning two, three, four, or five cars by 6.2%, 0.7%, 0.2%, and 0.002%, respectively. This latter figure is not statistically significant at the conventional levels.

3.5.2 Choice of emission category

I next turn to my second discrete choice model, which is of independent interest *and* is used to model for the selection of a vehicle's emissions class. This corresponds to the multinomial logit model specified in Equation (3.4) from Section 3.3. The results are outlined in Table 3.3. In my model, households purchasing a new vehicle choose one of five vehicle size classes, as defined by their emissions: 130 grams per kilometer (g/km) or less, 131 to 150 g/km, 151 to 200 g/km, 201 to 255 g/km, and 256 g/km and above.[28] Reported coefficients are interpreted relative to the lowest emitting category, which is omitted for identification purposes.

The first row of Table 3.3 reports coefficients for fuel prices for the 12 months leading up to the estimated month of purchase. These are intuitive, suggesting that higher fuel prices decrease the probability of purchasing cars in higher emitting categories (because these cars are less fuel efficient). Calculating the marginal effect for the average observation from the sample, a one percent increase in the price of fuel increases the probabilities of choosing the lowest emitting class by 0.48%, and decreases the probability of choosing the third and fourth classes by 0.35%and 0.36%. The marginal effects for the other two categories are statistically insignificant at the conventional levels.[29]

[28] In 2010, the last year of my study period, the first group was comprised of VED bands A, B, C and D, the second of E and F, the third of G. H and I, and the fourth and fifth would cover bands K, L and M.

[29] Additional statistically insignificant marginal effects predicted by the fuel price variable: for 131-150 g/km, 0.23 (standard error: 0.16); 255+ g/km: -0.007 (0.04).

Table 3.3. Multinomial logit regression results for vehicle emissions category chosen per vehicle (omitted category: <=130 grams per kilometer)

	(1)	(2)	(3)	(4)
VARIABLES	131-150	151-200	201-255	256+
log fuel price	-6.235***	-7.709***	-10.70***	-7.513***
	(1.092)	(1.033)	(1.564)	(2.587)
rural	-0.0245	-0.0237	0.219**	0.388**
	(0.0915)	(0.0872)	(0.108)	(0.166)
# other vehicles owned (up to year prior to purchase)	-0.145**	-0.104*	0.0434	0.360***
	(0.0635)	(0.0604)	(0.0739)	(0.106)
household income in 2nd quintile	-0.104	-0.346**	-0.208	-0.0835
	(0.175)	(0.167)	(0.237)	(0.447)
household income in 3rd quintile	0.0869	-0.124	0.189	-0.0121
	(0.170)	(0.162)	(0.224)	(0.430)
household income in 4th quintile	0.0344	-0.0614	0.172	0.305
	(0.170)	(0.162)	(0.223)	(0.413)
household income in 5th quintile	0.183	0.268	1.092***	1.491***
	(0.174)	(0.166)	(0.223)	(0.402)
Household owns home	0.0444	-0.0268	0.181	0.725**
	(0.128)	(0.122)	(0.168)	(0.339)
Year 2002	-0.183	-0.546**	-0.807***	-0.647*
	(0.272)	(0.263)	(0.291)	(0.386)
Year 2003	-0.732***	-1.213***	-1.364***	-1.341***
	(0.251)	(0.242)	(0.271)	(0.374)
Year 2004	-0.544**	-1.039***	-1.216***	-1.102***
	(0.249)	(0.240)	(0.266)	(0.360)
Year 2005	-0.556**	-0.891***	-1.068***	-1.031***
	(0.242)	(0.233)	(0.260)	(0.354)
Year 2006	-0.531**	-0.763***	-0.685**	-0.686*
	(0.251)	(0.239)	(0.282)	(0.400)
Year 2007	-0.695***	-1.125***	-1.556***	-1.348***
	(0.241)	(0.230)	(0.275)	(0.386)
Year 2008	-0.575*	-1.005***	-1.249***	-1.499***
	(0.294)	(0.280)	(0.367)	(0.578)
Year 2009	-1.204***	-1.783***	-2.000***	-2.228***
	(0.278)	(0.265)	(0.346)	(0.557)

Year 2010	-1.449***	-2.196***	-2.870***	-3.779***
	(0.302)	(0.291)	(0.496)	(1.102)
Family household, with adult child(ren)	-0.163	-0.448***	-0.821***	-0.923***
	(0.145)	(0.138)	(0.178)	(0.287)
Pensioner household	0.00184	-0.396**	-0.476**	-0.868**
	(0.167)	(0.160)	(0.208)	(0.353)
Other household	-0.132	-0.453***	-0.582***	-0.968***
	(0.103)	(0.0971)	(0.118)	(0.176)
2 persons in household with driver's license	0.0863	0.226**	0.406***	-0.0438
	(0.100)	(0.0962)	(0.127)	(0.208)
3 persons in household with driver's license	0.549***	0.369*	0.533**	-0.782*
	(0.196)	(0.189)	(0.242)	(0.425)
4+ persons in household with driver's license	0.562*	0.200	0.652*	-0.388
	(0.288)	(0.282)	(0.349)	(0.561)
Time from bus to rail station >14 mins.	-0.164*	-0.226**	-0.263**	-0.442**
	(0.0972)	(0.0927)	(0.115)	(0.177)
Walk time to rail station >26 mins	0.00650	0.0580	0.110	0.0923
	(0.104)	(0.0991)	(0.123)	(0.190)
Walk time to bus stop >6 mins.	0.00690	0.0672	0.286**	0.687***
	(0.109)	(0.104)	(0.125)	(0.176)
Socioeconomic group: clerical	-0.106	-0.345***	-0.770***	-0.895***
	(0.106)	(0.101)	(0.125)	(0.197)
Socioeconomic group: skilled manual	-0.0857	-0.257**	-0.417***	-0.385
	(0.133)	(0.127)	(0.155)	(0.241)
Other socioeconomic group / other manual	0.0200	-0.273	-0.734***	-0.768**
	(0.172)	(0.167)	(0.222)	(0.385)
Socioeconomic group: retired	0.0273	0.159	-0.0899	-0.0759
	(0.159)	(0.152)	(0.195)	(0.328)
Socioeconomic group: other economically inactive	0.267	-0.0219	0.405	0.900*
	(0.324)	(0.315)	(0.368)	(0.488)
NW & Merseyside	-0.0244	0.0471	0.303	-0.621*
	(0.194)	(0.188)	(0.257)	(0.373)
Yorkshire & Humberside	-0.00152	0.118	0.300	-0.598
	(0.204)	(0.197)	(0.269)	(0.396)
East Midlands	-0.157	0.0696	0.376	-0.596

	(0.204)	(0.196)	(0.266)	(0.392)
West Midlands	0.0655	0.268	0.262	-0.368
	(0.202)	(0.195)	(0.268)	(0.377)
Eastern Region	-0.107	0.180	0.542**	-0.186
	(0.201)	(0.193)	(0.260)	(0.359)
Greater London	-0.212	0.274	1.061***	0.457
	(0.238)	(0.226)	(0.288)	(0.393)
South East Region	-0.00508	0.299	0.683***	-0.192
	(0.192)	(0.185)	(0.250)	(0.347)
South West Region	-0.105	0.123	0.416	-0.0940
	(0.205)	(0.197)	(0.267)	(0.369)
Wales	-0.0416	0.0666	0.282	-0.894*
	(0.222)	(0.215)	(0.295)	(0.483)
Scotland	0.0779	0.137	0.359	-0.837**
	(0.197)	(0.192)	(0.263)	(0.408)
Constant	29.79***	37.53***	48.41***	33.08***
	(4.849)	(4.587)	(6.938)	(11.48)
Observations	10,694	10,694	10,694	10,694

Standard errors are reported in parentheses; the symbols ***, **, and * denote statistical significant at the 1%, 5%, and 10% levels, respectively.

Table 3.3 displays a number of other striking results Income is not a generally strong predictor of the choice of the vehicle's emissions class, except for households in the highest income quintile who exhibit a higher probability of purchasing the highest emitting vehicles. Higher emitting vehicles are also more likely to be purchased by homeowners. We also observe negative coefficients for all year fixed effects (relative to 2001, the omitted year), with overall increasing magnitudes over time. This finding is not surprising, since it corresponds to decreased demand for higher emitting vehicles over time, due in part to the VED which was increased multiple times during this period. Higher emitting vehicles are less likely to be purchased by households with adult children and pensioner households, relative to households with young children (the omitted household category)—a finding we might also expect, since higher passenger capacity cars tend to produce more emissions. With few exceptions, higher emitting vehicles are also more likely to be purchased by households with more registered drivers, households which are located further from bus transit (but also whose time from bus to rail is shorter), and managerial and other inactive socioeconomic groups (managerial being the omitted socioeconomic group). Among geographic regions, we observe a higher probability of owning vehicles in certain high emission categories in the Eastern region, South East region, and – interestingly - Greater London, while lower probabilities for higher emitting vehicles are observed for Wales, Scotland, and Northwest and Merseyside.

For rural households, the model reports positive coefficients for vehicles purchased from the two highest emitting categories. This finding is consistent with the higher average emissions previously noted among new vehicles of rural households in Figure 3.3. For the average observation in the sample, these coefficients translate into an increase in the probability of owning a vehicle emitting between 201 to 255 g/km and 256 g/km and above by 2.1% and 0.7 percent, respectively.[30]

3.5.3 Vehicle miles traveled

Having modeled the choice of a purchased vehicle's emissions category, I next use these results to model the demand for travel, utilizing the Conditional Expectation Correction Method, of the form specified in Equation (3.7). These results are reported in Table 3.4.[31] Column (1) begins with a specification using the full sample. The first row,

---

[30] Additional, statistically insignificant marginal effects predicted by the rural indicator: for 130 g/km and below, -0.001 (0.005); 131-150 g/km: -0.009 (0.011).

[31] In the Table A.2 of Appendix 2, I include analogous results which do not control for self-selection. The results are similar.

with a coefficient of -0.45, indicates that drivers are generally inelastic, with respect to the price per mile driven. This finding is generally consistent with previous findings of the literature. It reinforces the idea that using a tax, either directly on mileage or indirectly on transportation fuel, as a policy instrument to reduce externalities generated by transportation is unlikely to be very effective (Parry and Small, 2005; Li et al., 2014; Small and Van Dender, 2007). It also has implications for the rebound effect, by suggesting that negative shocks to the price of fuel are unlikely to cause major changes in driving behavior.

Other findings from the model are generally intuitive. Having a diesel car is associated with more driving, as is a car with higher engine capacity or in a larger size category (relative to small cars, the omitted category). One exception is the coefficient for vehicle age, as older vehicles (i.e. one to two years in age, versus less than one year) are driven more, whereas it is generally assumed that cars are driven most in their first year of life. The difference in this model, though, is relatively small, at less than five percent. Additional miles driven are also associated with higher income households with children, more drivers, fewer total cars, and with managerial/professional duties (the omitted socioeconomic category). Geographically, fewer miles are driven in Greater London. I also observe more driving among households located further from rail stations, who likely have fewer travel alternatives. The coefficient of main interest, rural status, increases the number of miles driven by 12.8 percent.

Table 3.4. Travel demand regression results, including Conditional Expectation Correction Method

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log price (pence) per 100km	-0.451***	-0.644***	-0.540***	-0.435***	-0.487***	-0.543***	-0.518***
	(0.0823)	(0.0618)	(0.128)	(0.113)	(0.0937)	(0.133)	(0.109)
log price (pence) per 100km x (130-149g/km emissions)					0.0215*	0.0474*	0.0321**
					(0.0122)	(0.0243)	(0.0152)
log price (pence) per 100km x (150-199g/km emissions)					0.0115	0.0513**	0.0394***
					(0.0127)	(0.0252)	(0.0138)
log price (pence) per 100km x (200-249g/km emissions)					0.0209	0.0540*	0.0572***
					(0.0163)	(0.0294)	(0.0207)
log price (pence) per 100km x (>=250g/km emissions)					-0.0223	0.0266	0.0149
					(0.0251)	(0.0369)	(0.0318)
rural	0.128***	0.127***	0.134***	0.126***	0.129***	0.137***	0.121***
	(0.0131)	(0.0131)	(0.0221)	(0.0216)	(0.0131)	(0.0189)	(0.0184)
diesel	0.201***	0.165***	0.205***	0.222***	0.201***	0.201***	0.212***
	(0.0214)	(0.0193)	(0.0391)	(0.0298)	(0.0215)	(0.0354)	(0.0267)
log engine capacity	0.373***	0.414***	0.400***	0.318***	0.376***	0.429***	0.294***
	(0.0475)	(0.0458)	(0.0719)	(0.0730)	(0.0476)	(0.0679)	(0.0675)
age of car 1-2 years	0.0497***	0.0474***	0.0661***	0.0817***	0.0492***	0.0413**	0.0487***
	(0.0117)	(0.0115)	(0.0182)	(0.0184)	(0.0119)	(0.0164)	(0.0165)
small/medium car	0.0880***	0.0889***	0.110***	0.0754**	0.0868***	0.0957***	0.0707**

	(0.0194)	(0.0193)	(0.0291)	(0.0324)	(0.0194)	(0.0265)	(0.0290)
medium car	0.231***	0.229***	0.262***	0.238***	0.230***	0.241***	0.211***
	(0.0223)	(0.0223)	(0.0356)	(0.0353)	(0.0224)	(0.0318)	(0.0320)
large car	0.323***	0.326***	0.374***	0.334***	0.320***	0.322***	0.302***
	(0.0270)	(0.0270)	(0.0415)	(0.0429)	(0.0270)	(0.0380)	(0.0387)
land rover / jeep	0.144***	0.153***	0.173***	0.155***	0.145***	0.155***	0.124**
	(0.0362)	(0.0361)	(0.0605)	(0.0593)	(0.0362)	(0.0516)	(0.0505)
household income in 2nd quintile	-0.00220	-0.00280	-0.0266	0.0136	-0.00343	-0.0329	0.0316
	(0.0277)	(0.0277)	(0.0426)	(0.0426)	(0.0281)	(0.0398)	(0.0393)
household income in 3rd quintile	0.0805***	0.0789***	0.104**	0.0545	0.0792***	0.0845**	0.0791**
	(0.0261)	(0.0262)	(0.0406)	(0.0405)	(0.0264)	(0.0380)	(0.0369)
household income in 4th quintile	0.149***	0.148***	0.190***	0.110***	0.148***	0.171***	0.129***
	(0.0256)	(0.0257)	(0.0403)	(0.0406)	(0.0258)	(0.0364)	(0.0370)
household income in 5th quintile	0.193***	0.194***	0.229***	0.155***	0.193***	0.217***	0.166***
	(0.0259)	(0.0259)	(0.0413)	(0.0409)	(0.0261)	(0.0380)	(0.0367)
Family household, with adult child(ren)	0.0223	0.0189	0.0602*	0.0132	0.0216	0.0306	0.0248
	(0.0212)	(0.0212)	(0.0342)	(0.0348)	(0.0214)	(0.0313)	(0.0293)
Pensioner household	-0.206***	-0.213***	-0.179***	-0.235***	-0.208***	-0.198***	-0.197***
	(0.0251)	(0.0251)	(0.0402)	(0.0398)	(0.0255)	(0.0371)	(0.0351)
Other household	-3.97e-05	-0.00290	0.0303	-0.0120	-0.00157	0.0250	-0.0157
	(0.0147)	(0.0146)	(0.0231)	(0.0240)	(0.0151)	(0.0205)	(0.0216)
NW & Merseyside	-0.0104	-0.00849	-0.0928*	-0.00603	-0.00996	-0.0655	0.0341
	(0.0291)	(0.0291)	(0.0477)	(0.0478)	(0.0291)	(0.0424)	(0.0404)
Yorkshire & Humberside	0.00188	0.00353	-0.113**	0.0294	0.00150	-0.0802*	0.0670
	(0.0305)	(0.0305)	(0.0494)	(0.0505)	(0.0306)	(0.0444)	(0.0425)
East Midlands	0.00596	0.00847	0.00935	-0.0248	0.00770	0.0125	0.00213
	(0.0313)	(0.0313)	(0.0504)	(0.0507)	(0.0314)	(0.0451)	(0.0436)
West Midlands	-0.0118	-0.00842	-0.0609	0.00514	-0.0103	-0.0544	0.0256

	(0.0301)	(0.0301)	(0.0486)	(0.0492)	(0.0303)	(0.0433)	(0.0425)
Eastern Region	-0.0214	-0.0179	-0.108**	0.00839	-0.0194	-0.0739*	0.0247
	(0.0299)	(0.0299)	(0.0489)	(0.0493)	(0.0300)	(0.0435)	(0.0415)
Greater London	-0.241***	-0.234***	-0.296***	-0.254***	-0.239***	-0.270***	-0.228***
	(0.0361)	(0.0361)	(0.0538)	(0.0566)	(0.0364)	(0.0509)	(0.0520)
South East Region	-0.0350	-0.0316	-0.0591	-0.0547	-0.0340	-0.0530	-0.0228
	(0.0286)	(0.0286)	(0.0467)	(0.0466)	(0.0288)	(0.0421)	(0.0396)
South West Region	-0.00435	0.000331	-0.0586	0.0162	-0.00182	-0.0585	0.0432
	(0.0309)	(0.0309)	(0.0494)	(0.0511)	(0.0310)	(0.0446)	(0.0432)
Wales	0.0289	0.0321	-0.0460	0.0613	0.0296	-0.0290	0.0786
	(0.0350)	(0.0350)	(0.0556)	(0.0560)	(0.0350)	(0.0506)	(0.0487)
Scotland	-0.0432	-0.0447	-0.119**	-0.0603	-0.0437	-0.0767*	-0.0181
	(0.0300)	(0.0300)	(0.0494)	(0.0486)	(0.0300)	(0.0442)	(0.0415)
Time from bus to rail station >14 mins.	-0.0301**	-0.0321**	-0.0553**	-0.00980	-0.0306**	-0.0510**	-0.00864
	(0.0141)	(0.0141)	(0.0226)	(0.0226)	(0.0141)	(0.0199)	(0.0201)
Walk time to rail station >26 mins	0.0298**	0.0316**	0.0395	0.0178	0.0309**	0.0332	0.0265
	(0.0150)	(0.0150)	(0.0241)	(0.0240)	(0.0150)	(0.0212)	(0.0214)
Walk time to bus stop >6 mins.	-0.00747	-0.00636	-0.0578**	-0.00188	-0.00648	-0.00758	-0.0116
	(0.0155)	(0.0155)	(0.0255)	(0.0252)	(0.0156)	(0.0224)	(0.0219)
Household owns home	-0.0191	-0.0158	0.00178	-0.0314	-0.0198	-0.0218	-0.00864
	(0.0220)	(0.0219)	(0.0329)	(0.0311)	(0.0220)	(0.0335)	(0.0292)
Socioeconomic group: clerical	-0.0448***	-0.0475***	-0.0299	-0.0576**	-0.0450***	-0.0385*	-0.0429*
	(0.0153)	(0.0153)	(0.0249)	(0.0252)	(0.0154)	(0.0218)	(0.0220)
Socioeconomic group: skilled manual	-0.144***	-0.143***	-0.103***	-0.148***	-0.143***	-0.108***	-0.175***
	(0.0199)	(0.0199)	(0.0303)	(0.0320)	(0.0199)	(0.0270)	(0.0298)
Other socioeconomic group / other manual	-0.126***	-0.128***	-0.117***	-0.181***	-0.128***	-0.0892***	-0.170***

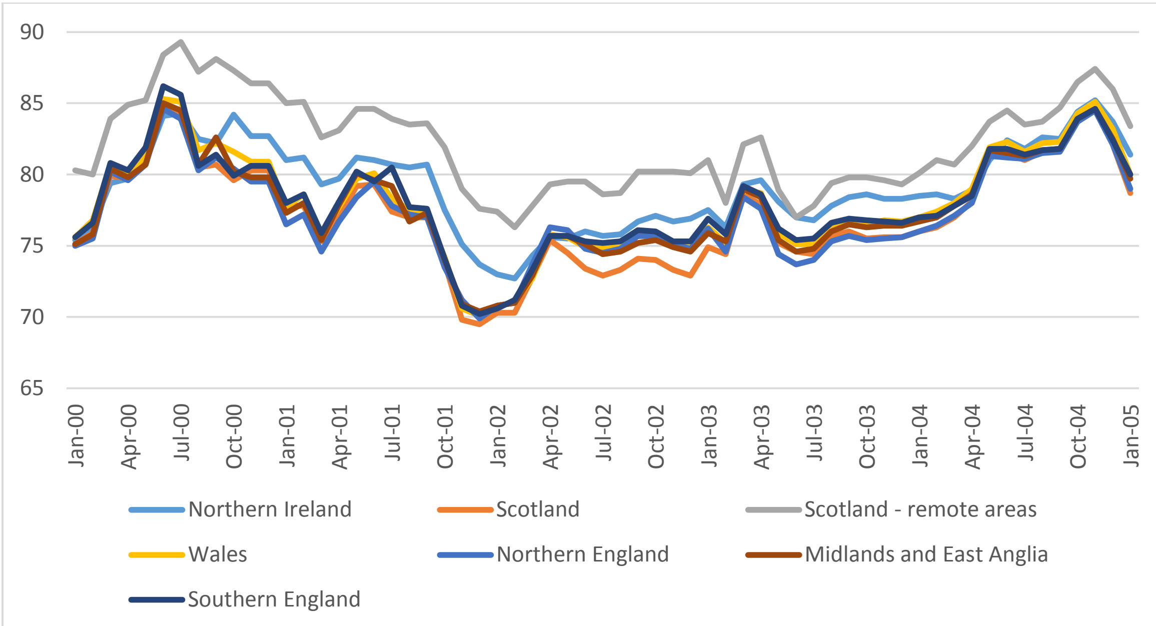
	(0.0249)	(0.0251)	(0.0398)	(0.0432)	(0.0251)	(0.0340)	(0.0375)
Socioeconomic group: retired	-0.276***	-0.274***	-0.296***	-0.335***	-0.274***	-0.258***	-0.302***
	(0.0238)	(0.0237)	(0.0373)	(0.0379)	(0.0240)	(0.0360)	(0.0319)
Socioeconomic group: other economically inactive	-0.226***	-0.224***	-0.181**	-0.317***	-0.225***	-0.161**	-0.289***
	(0.0498)	(0.0499)	(0.0720)	(0.0747)	(0.0500)	(0.0692)	(0.0718)
2 persons in household with driver's license	-0.0182	-0.0187	-0.0143	-0.00182	-0.0167	-0.0245	-0.0164
	(0.0147)	(0.0147)	(0.0227)	(0.0236)	(0.0148)	(0.0204)	(0.0212)
3 persons in household with driver's license	0.00962	0.00652	0.00169	0.0286	0.00792	0.0131	0.00757
	(0.0266)	(0.0266)	(0.0433)	(0.0431)	(0.0266)	(0.0390)	(0.0366)
4+ persons in household with driver's license	0.00406	0.000606	0.0485	0.0436	0.000188	0.0124	0.00252
	(0.0363)	(0.0365)	(0.0667)	(0.0639)	(0.0366)	(0.0559)	(0.0489)
Survey year: 2002	0.101*				0.101*		
	(0.0579)				(0.0579)		
Survey year: 2003	0.0159				0.0137		
	(0.0471)				(0.0473)		
Survey year: 2004	0.0401				0.0380		
	(0.0497)				(0.0501)		
Survey year: 2005	0.0200				0.0201		
	(0.0508)				(0.0513)		
Survey year: 2006	-0.00845				-0.00661		
	(0.0507)				(0.0513)		
Survey year: 2007	0.00759				0.00850		
	(0.0511)				(0.0518)		
Survey year: 2008	-0.0212				-0.0161		
	(0.0535)				(0.0558)		
Survey year: 2009	-0.0662				-0.0619		
	(0.0530)				(0.0553)		

Survey year: 2010	-0.0594				-0.0451		
	(0.0613)				(0.0667)		
DM correction term: <130	0.0437***	0.0746***	0.0233	0.0792***	0.0173	-0.0669	-0.00609
	(0.0153)	(0.0136)	(0.0217)	(0.0243)	(0.0250)	(0.0442)	(0.0361)
DM correction term: 130-149	0.0540***	0.0821***	0.0309	0.0797***	0.109**	0.0293	0.101
	(0.0207)	(0.0197)	(0.0307)	(0.0304)	(0.0498)	(0.0546)	(0.0650)
DM correction term: 150-199	0.0257	0.0446	-0.0548	0.0813	0.0246	-0.00355	0.184**
	(0.0428)	(0.0423)	(0.0705)	(0.0576)	(0.0683)	(0.0903)	(0.0737)
DM correction term: 200-249	-0.00171	-0.0182	-0.0485*	0.0209	0.0244	-0.00871	0.0845*
	(0.0166)	(0.0158)	(0.0254)	(0.0243)	(0.0326)	(0.0476)	(0.0461)
DM correction term: 250+	-0.00325	-0.0243	-0.0363	0.0200	-0.0731*	-0.0663	-0.0297
	(0.0162)	(0.0151)	(0.0251)	(0.0275)	(0.0407)	(0.0489)	(0.0559)
Constant	9.012***	9.974***	9.301***	9.364***	9.152***	8.877***	9.876***
	(0.490)	(0.419)	(0.741)	(0.707)	(0.550)	(0.759)	(0.686)
Observations	10,469	10,469	5,128	5,527	10,469	5,043	5,426
R-squared	0.278	0.276	0.269	0.249	0.278	0.292	0.270
Sample period	2002-2010	2002-2010	2002-2005	2006-2010	2002-2010	2002-2005	2006-2010

Standard errors in parentheses, and are clustered at the household level; the symbols ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

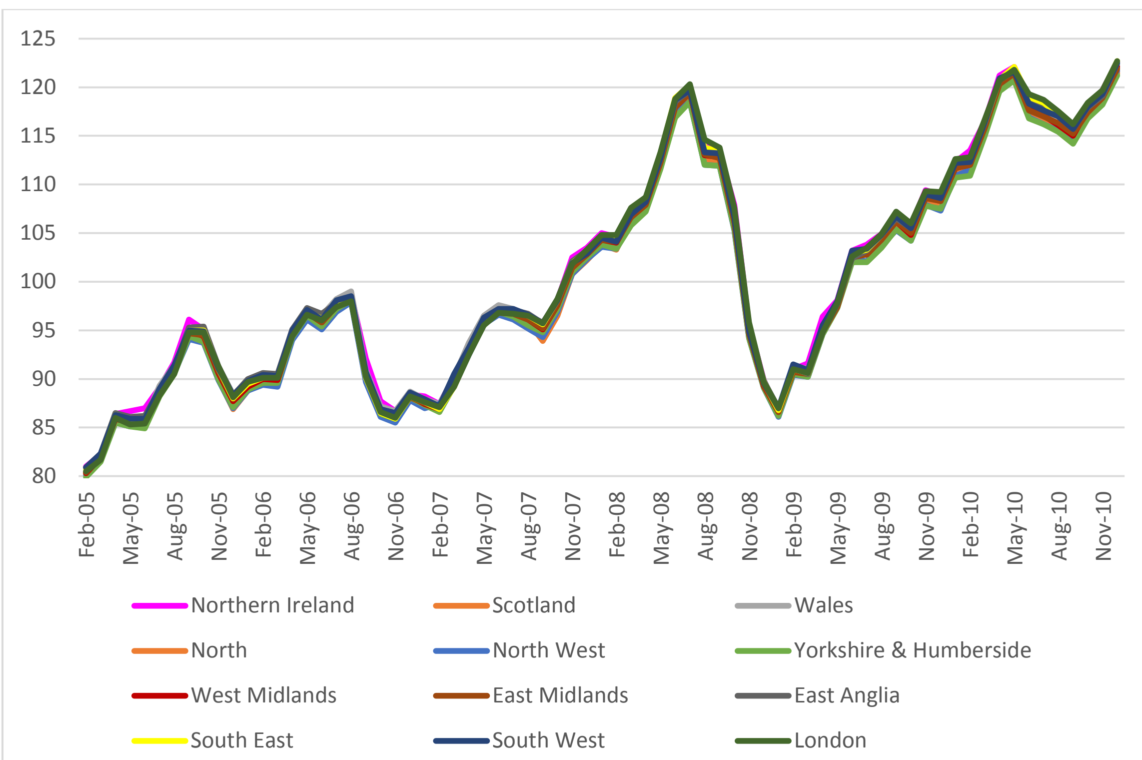
In Column (2), I explore the sensitivity of the coefficients, particularly for the price per mile, to the use of time period fixed effects. The reason is that, controlling for the time period of being surveyed, the variation in fuel prices – which is used to calculate the price per mile - is rather small, since my data for fuel prices is aggregated at the level of the GOR on a monthly basis, and prices at these levels tend not to differ that much from one another. This is illustrated in Figures 3.4 and 3.5, which plot the monthly prices for unleaded-95 fuel, from 2000 to 2005, and 2005 to 2010 respectively. I use 2005 as a key date because the geographic regions (used to aggregate fuel prices) were changed in February, 2005. These figures show not only that prices vary little between regions, but that they also move with one another, over time (similar trends are observed for diesel prices). Column (2) of Table 3.4 illustrates that when year fixed effects are omitted, the price elasticity is higher. This is to be expected: since the price per mile is a function of fuel prices, and the variation across observations for fuel price is limited, year fixed effects absorb some of the variation observed from fuel prices. Interestingly, the coefficients for most all the other variables are fairly robust to the omission of year fixed effects – including the rural status indicator.

Figure 3.4. Price trends for Unleaded 95 fuel, by Government Office Region: January, 2000-January, 2005



Source: The Automobile Association, Fuel Price Reports

Figure 3.5. Price trends for Unleaded-95 fuel, by Government Office Region: Feb. 2005-Dec. 2010

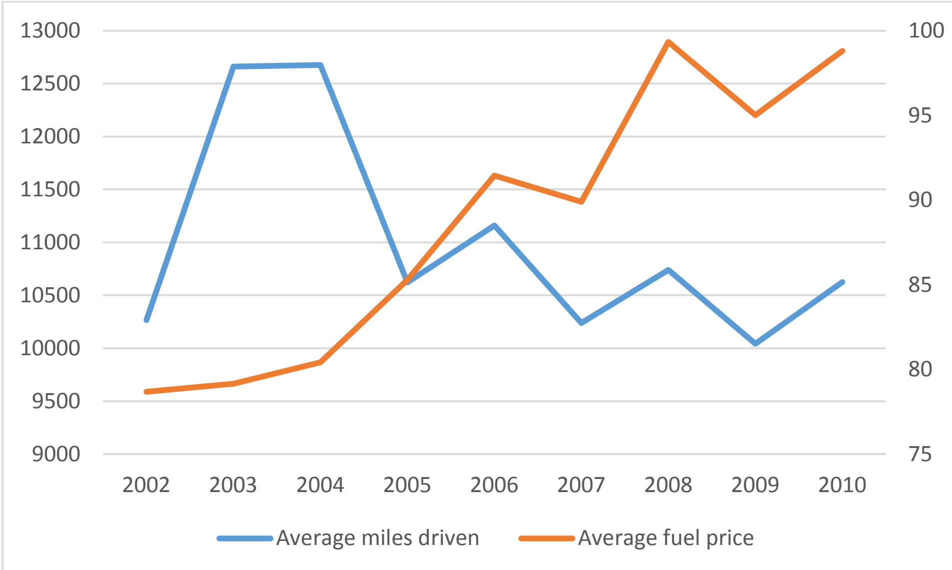


Source: The Automobile Association, Fuel Price Reports

In columns (3) and (4), I assess the sensitivity of the coefficients to the omission of time dummy variables, by splitting the sample into two separate time periods. Column (3) considers the results using the first four years of the sample, 2002 to 2005, while columns (4) consider the last five years, 2006 to 2010. For the most part, results tend to be stable over time. The only exception is the earlier noted coefficient for the price per mile: from 2006 to 2010, drivers appear to be about 20% less responsive to the price per mile as they were between 2002 and 2005, although a Chow Test fails to reject the hypothesis that these coefficients are equal. This is an interesting finding; as illustrated from Figure 3.5, 2006 to 2010 was a more unstable period, with fuel prices both peaking and plummeting. The low price elasticity could be a result of consumer adaptation to higher fuel prices, perhaps from switching to more fuel efficient cars during this time period. However, this seems unlikely, since I already control for vehicle selection by including the Conditional Expectation Correction terms.

The difference in price elasticities observed between the two time periods more likely reflects the difference in economic conditions between the two periods. Around the same time that fuel prices began to decline, following the financial crisis of 2008, consumers also started driving less. This is shown in Figure 3.6, which plots the average number of annual miles driven in my data against the mean 12-month averaged fuel price calculated for each vehicle. It is clear that, from 2006 to 2010, the negative relationship between fuel prices and miles driven is much weaker. These coinciding economic trends causing both consumer activity and fuel prices to move in similar directions during the 2006 to 2010 period arguably leads to a weaker negative relationship between the number of miles driven and the price of fuel – and, by extension, the price per mile.

Figure 3.6. Average miles driven per year, and average fuel price



Source: UK National Travel Survey
Note: averages are calculated for each year, using respondents of cars age 2 or less who completed the survey during one of the months of that year; the fuel price associated with each respondent is calculated by taking the average fuel price observed for that household's region for the preceding 12 months (the same period over which the vehicle miles were driven).

I next consider the heterogeneity of price elasticities on the basis of emissions class, using the same groups identified as choice outcomes for the multinomial logit model from Section 3.5.2. These results, reported in column (5), include interaction terms between the price per mile and four of the five emissions categories, and once again include all years of the sample. The lowest emitting category - cars emitting fewer than 130 grams per kilometer - is omitted, for identification purposes. This produces the equivalent of equation (3.7) in Section 3.3, where the coefficient for the price per mile is permitted to vary between vehicle classes. These results suggest that the fuel price elasticity, overall, is fairly constant across these categories - with the minor exception of the coefficient for the term interacting price with the second highest class of emitting vehicles that is four percent lower than other vehicle classes. Columns (6) and (7) serve as the analog to columns (3) and (4), by considering the results separately for two time periods, without year fixed effects. These columns suggest that higher emitting vehicles may actually exhibit lower price elasticity than lower emitting vehicles, although the differences are quite small – no more than just over 10%, relative to the lowest emitting group of cars. This may reflect the underlying heterogeneity of the drivers themselves, as those with larger vehicles may use their vehicle more for occupational purposes, and may be less sensitive to the volatility of the fuel prices observed during this period.

3.6 Discussion

Turning to my final research question, I consider the implications of these findings on carbon footprints. Rural households own more, higher emitting vehicles, with which they drive more miles, than urban households. Together, these findings suggest that these households produce more emissions with their vehicles. With this information, I do a series of back-of-the-envelope calculations, to estimate just how much more. These estimates are summarized in Table 3.5. The UK and Scotland Censuses together classify 5.03 million households, 19.5 percent of the total, as rural in origin (this compares to 20.6 percent of all households, in the NTS). I estimate the mean number of cars owned by each of these segments, urban and rural, for the year 2010 - by comparing the mean of each group in my data – to be 1.12 and 1.55 vehicles, respectively. This is, of course, consistent with earlier findings of this chapter, which document higher vehicle ownership rates among rural households.

Table 3.5. Comparison of carbon emissions between urban and rural households

	Urban	**Rural**	**Source**
Total households (millions)	20.71	5.03	UK Office for National Statistics, 2011 Census; Scotland's Census, 2011*
Vehicles per household (2010)	0.99	1.35	UK National Travel Survey
Emissions per vehicle (grams per mile, 2010)	244.00	248.60	UK National Travel Survey**
Average annual mileage per vehicle (2010)	8084.41	9188.36	UK Department of Transport, National Travel Survey Statistics***
Total emissions per household (tons)	2.15	3.40	
Share of total UK vehicle emissions	72.3%	27.7%	

*To maintain consistency with the UK Census' classification of "rural", households in the Scotland census were identified as rural if they resided in "accessible towns", "remote towns", "accessible rural", and "remote rural" areas, as these four groups are classified as having populations below 10,000

**Estimated from regression of emissions on time trend for year of purchase and rural status

***The mean mileage for all vehicles in 2010, provided by such, is used as a base. Information for the proportion of urban and rural households within the total population is then used with the rural coefficient calculated from the mileage regression to estimate annual mileage for each group.

Next, I estimate the average emissions per vehicle for each group, by running a simple regression of vehicle emissions on a time trend, corresponding to the estimated year of purchase, and rural status. Note that emissions, previously reported in this chapter in terms of grams per kilometer, are reported here in grams per mile, for ease of multiplication with the next term, average mileage.[32] To calculate this, I use the official mean estimate for vehicle mileage provided by the UK Department of Transportation, National Travel Survey Statistics, which for 2010 is 8,300 miles. Using the exponential of the coefficient for rural status from Column (1) of Table 3.4, combined with the known household count of urban and rural households, allows me to estimate, algebraically, the average mileage for each group, 8,101 and 9,216 miles respectively.[33] The product of the latter three averages - vehicles per household, emissions per mile per vehicle, and the number of miles – is the annual emissions produced by urban and rural households to be 2.15 and 3.40 tons, respectively. By these calculations, in 2010 rural households emitted 58% additional emissions with their vehicles. This translates to 19.5 percent of households being responsible for just over one quarter, 27.7 percent, of vehicular emissions in the UK.

3.7 Conclusions

Given the number of externalities associated with transportation - most notably, carbon dioxide emissions in the current political climate – it is important for policymakers to consider the geographic distribution of these externalities, when assessing options for how to address them. This chapter, in particular, focuses on the differences in driving behavior between rural and urban households. Using travel survey microdata from the UK, it specifically considers how differences in vehicle ownership,

[32] Indexing the first year, 2002, with a value of one, this result produces the equation Emissions = 178.85 - 2.74(Time Trend) + 2.86(Rural). As the NTS reports emissions in grams per kilometer, the estimates are multiplied by 1.61 to convert the denominator to miles.

[33] The equation solved for here takes the form: 8,300 = (X * 0.815) + ((X* $e^{0.128}$)*0.185); where X equals the number of urban miles, X* $e^{0.128}$ the number of rural miles, and 0.815 and 0.185 the shares of urban and rural households, respectively.

fuel efficiency, and travel, and their ultimate impact on carbon dioxide emissions. Among rural households, I find higher amounts of vehicle ownership, illustrated by both a direct comparison of summary statistics, and a discrete choice model for the number of cars chosen by a household to own. This is consistent with expectations, since rural regions provide fewer options for alternative transportation. Not only do they own more vehicle, but on average these vehicles are higher emitting, although this difference is relatively small. Modeling the demand for travel, using a framework which accounts for the selection of a particular vehicle type, based on the vehicle's emissions, I find that rural households drive their vehicles 10.7 percent more than urban households. Taken together, I estimate a carbon footprint of 3.40 tons per rural household, which is 58 percent higher than those of urban households. Collectively, they account for 27.7 percent of carbon dioxide emissions from vehicular transportation.

A few limitations are worth noting, from this exercise. First, the lack of specificity of data from the UK National Travel Survey, concerning the vehicle make and model, required me to use broader categories, for modeling vehicle selection. Such an exercise could be improved and made more informative with such data. Second, the survey provided little information on the specific geography, or individual fuel prices, faced by households. Such information might prove useful for the purposes of increasing price variation in the data.

Lastly, the calculated gap in emissions generated by transportation more generally, between rural and urban households, are arguably somewhat overstated, once one accounts for the emissions generated from public transportation, primarily utilized by urban households, as well as the reduction in fuel efficiency that arises from driving at slower speeds in more densely populated areas (Emrath and Liu, 2008; Nechyba and Walsh, 2004).

Chapter 4. Methamphetamine lab discoveries and property values: a hedonic price analysis

4.1 Introduction

Crime prevention is a good which consumers place a value on, but how much? Like many amenities – such as environmental goods, and reduced health risks - it is a good for which there is no direct market. While there are other markets from which such value might be gauged (e.g. home security systems, self-defense equipment and courses, etc.), it is impractical to collect the data needed to form demand functions and estimate the willingness to pay for crime reductions. Nonmarket valuation presents an alternative method. One possible nonmarket approach for doing so is to rely on revealed preferences, whereby the value of a nonmarket good is estimated indirectly from the observed behavior of consumers in related markets. Revealed preference studies are often conducted in the form of the hedonic price regressions. Consider a market with differentiated goods in which buyers and sellers are matched, forming a price schedule where the price of a good is a function of its attributes. When such attributes include amenities concerning environmental quality or well-being, the marginal value of these amenities can be implicitly inferred from the coefficients of variables controlling for them. Rosen's (1974) seminal work formally illustrates that non-marginal welfare estimates can be obtained from the hedonic price model without any information about the underlying preferences of the consumers involved.

Hedonic price models have been applied to various products, including wage outcomes in the labor market and vehicle purchases. It is most commonly applied, however, to real estate property values, to estimate the value of (dis)amenities such as reductions in cancer risk (Davis, 2004), bacterial contamination of waterways (Leggett and Bockstael, 2000), air pollution (Chay and Greenstone, 2005), hazardous waste site cleanup efforts (McCluskey and Rausser, 2003; Greenstone and Gallagher, 2008), airport noise (Pope, 2008a), primary school education quality (Black, 1999), and leaks from underground storage tanks (Guignet, 2013).

Depending on the amenity of interest, however, hedonic property models can be difficult to employ. A major concern often underlying such models is omitted variable bias. Amenities potentially affecting a home's value are often correlated with additional characteristics which are harder to measure and missing from one's dataset, biasing point estimates for such amenities as a result. One cannot simply consider the value of school quality, for example, by comparing the value of homes in neighborhoods with access to different quality programs, since there tend to be preexisting differences that are hard to measure (such as neighborhood safety or additional neighborhood amenities).

When omitted variables are a concern, the ideal approach is to find a situation where a shock exogenously changes the level of an amenity relative to comparable homes in the same market. These situations are known as natural experiments, or quasi-experiments, and ensure that a property's exposure to an amenity is uncorrelated with any other variables which factor into its price. In practice, these situations are challenging to find. Several works, however, offer exemplary applications. Davis (2004), for example, considers the case of Churchill County, Nevada, which experienced an unanticipated outbreak of pediatric leukemia over a short period of time. Comparing changes in home values of the affected county to those of a neighboring control county, which prior to the outbreak had observed comparable sales trends, he estimates the value of health risk from the decreased property values observed in the affected county. Chay and Greenstone (2005) study the effect of air pollution reductions on property values across the US, using a quasi-experimental regression discontinuity design resulting from federal regulations from the 1970s which imposed additional pressure on US counties with air pollution levels above "attainment status".

In the case of crime, it is difficult to identify natural experiments which allow researchers to isolate its effect on home values.[34] Crime is not a purely random phenomenon, geographically speaking, as its prevalence tends to covary with other neighborhood qualities. It may be possible, however, to identify homes or neighborhoods affected by local randomness, such that while the assignment of the treatment may be nonrandom, the precise location of the treatment itself is arguably more so. This approach is employed by Linden and Rockoff (2008) and Pope (2008), in examining the effect of sex offenders, whose entry into a neighborhood is disclosed in a public register. Their findings suggest that homes located closer to sex offenders observe lower property values.

The case of methamphetamine laboratories (henceforth referred to as meth labs) is of particular interest to researchers and homebuyers alike. Methamphetamines are an illegal substance consumed for recreational drug use. Its production process is simple, enabling individuals to produce the substance with minimal input requirements. As a result, small scale operations which occur in residential homes are a common practice. In

[34] Some work tries to estimate the effect of crime on property values by other means. Pope and Pope (2012), for example, use a nationwide, zip-code level dataset to consider the effect from the precipitous drop in crime rates during the 1990s, using national crime trends as an instrumental variable to predict local rates. They admit, however, that "it is difficult to fully conclude that [their] IV estimates … are identifying the causal impact of crime on property values."

the past decade, there has been a surge in meth labs in many parts of the country, most of which – by one estimate – have yet to be discovered.[35] Like sex offender registers, the locations of discovered labs are publicly disclosed. Unlike homes engaged in other types of illicit activity, such as "crack houses" or those hosting prostitution, though, the location of these operations are often unknown to neighbors prior to their being raided by law enforcement. For this reason, discovery is an excellent "quasi-experiment" to study.

There are two reasons we would expect the discovery of meth labs to affect property values. First, they are a visible crime scene. Their presence on a register could serve as a signal for the presence of other, undiscovered meth labs in the neighborhood, and/or other crime in a neighborhood, more generally. Second, they are hazardous waste sites, which could pose a risk to public health. There is work which documents the connection between hazardous waste sites and property values (Kiel and Williams, 2007; Greenstone and Gallagher, 2008; Gamper-Rabindrana and Timmins, 2013)

Empirical work quantifying the effect of discovered labs on property values is now underway. Congdon-Hohman (2013), use a difference-in-difference (DID) approach using a sample of home sales in Summit County, Ohio. Briefly, a difference-in-difference approach focuses on a treatment, and seeks to identify the effect of this treatment (here, the discovery of a meth lab) by comparing the value of an outcome variable (here, property prices) before and after the treatment, using a sample of control units (homes that did not receive the "meth lab treatment") to filter out the possible influence of economy-wide trends that affect prices. For this approach to work, there are

[35] While the proportion of undiscovered meth labs is of course unknown, some estimates include 75% (Martin, 2014) and 90% (Easter, 2010).

specific data requirements that must be satisfied, and units must be assigned exogenously to the treatment or control groups. In Congdon-Hohman (2013), homes sold after the discovery of a meth lab are compared with homes sold both within similar ranges of proximity before such discoveries were made and in neighborhoods that are similar but more distant from the discoveries. He estimates a decrease in sales prices as high as 10 to 19%.

This chapter builds on Congdon-Hohman's analysis of meth labs, by providing several extensions. First, in his examination, the author relies on the national register provided by the federal Drug Enforcement Agency (DEA) to identify the location of discovered meth labs in his county. In addition to the DEA's register, localities at the state, county, and city level often provide their own registers for reference as well. In the case of Summit County, there are two additional listings: one provided by the County Sheriff, and one provided by the police department of its main city, Akron. Comparing the entries of the three registers with one another reveals that a sizable fraction of discovered labs are not included in the DEA's register. In addition, the Summit County registers is particularly useful for the detailed descriptions provided about many of its discoveries. In this dissertation chapter, I revisit the housing market in Summit County, Ohio, using these additional registers to differentiate - albeit crudely - between larger scale discoveries (e.g. a meth lab) from minor sites (e.g. a site where byproduct waste was identified).

Second, using these registers, this chapter explores another important factor: the sensitivity of the treatment effect to the discovery being publicly disclosed in a register. Previous results assume the availability of perfect information to homebuyers following the discovery of each meth lab. However, there are many labs which were documented in one or more registers, but were discovered before the registers themselves were published. If meth labs are unknown to their neighborhoods before their discovery, it is reasonable to expect that, in the absence of a register, their discoveries may remain unknown to homebuyers, who may be unfamiliar with a neighborhood's history. This chapter accounts for this, by differentiating between homes which sold at a time that a previous discovery had not been documented, and those sold at a time that nearby discoveries were documented.

Third, this chapter extends the identification strategy by additionally using repeat sales to estimate the treatment effect. To the extent that unobserved, potentially confounding variables are time invariant, using repeat sales "differences" them out of the regression equation altogether, and thus eliminates the associated biases.

The remainder of this chapter is organized as follows. Section 4.2 provides more background on the hedonic price model, omitted variables in such a context, and methamphetamine labs. Section 4.3 discusses the data, and Section 4.4 the empirical methodology. The results are presented in Section 4.5. Section 4.6 concludes.

4.2 Background

4.2.1 The hedonic price model

In the absence of markets where certain environmental amenities or health risks are traded, one must resort to nonmarket valuation methods to place a value on them. One approach (revealed preferences) seeks market goods that incorporate such amenities/risks into the price. The hedonic price model is one such commonly used method in this literature. Rosen (1974) provides the theoretical underpinnings of this model in its application to the housing market, assuming a sufficiently large quantity of homes and differentiation among attributes, by outlining the sorting process through which households with differing preferences bid for these properties. An equilibrium ensues from this process, with a price schedule corresponding to each attribute of interest. The model begins with a household which, in choosing a housing unit, chooses to maximize utility subject to its budget constraint:[36]

$$(4.1)\ \ \max_{z,\boldsymbol{q}}\ u(z, \boldsymbol{q};\ \delta)\ \text{ subject to }\ y = z + P(\boldsymbol{q}).$$

Here, z represents a numeraire good, y the household's income, δ household-specific preferences, $\boldsymbol{q}$ a K-dimensional vector of housing unit characteristics, each element of which is chosen by households in this problem, and $P(\boldsymbol{q})$ the price of the housing unit. Setting up the Lagrangian and solving for the first order conditions ultimately gives us the following K conditions:

$$(4.2)\ \frac{\partial u(q, y-P(\boldsymbol{q});\ \delta)/\partial q_k}{\partial u(q, y-P(\boldsymbol{q});\ \delta)/\partial z} = \partial P(\boldsymbol{q})/\ \partial q_k,\ k = 1, \ldots, K.$$

For each *kth* housing attribute, the marginal rate of substitution between the attribute and the numeraire good (i.e., the slope of the indifference curve for the two goods) equates to the marginal change in the price of the housing unit with respect to that attribute, i.e. the slope of the hedonic price function.

[36] For additional information on the Rosen model, see Bockstael and McConnell (2007).

One can use this model to derive, for a fixed level of utility, income, and all other housing attributes of $\boldsymbol{q}$, a household's bid function for q_k, i.e. the maximum amount that it would be willing to pay for a given amount of q_k. The hedonic price function itself, which for any one household is taken as given, results from the equilibrium process assigning each housing unit to the highest bidder. As households vary in terms of preferences and income, an allocation process occurs whereby no household willingly outbids other households for any other unit.[37] The hedonic price function that emerges from this process serves as an envelope of the bid functions of participating households. As such, for any given attribute q_k, the hedonic price function $P(q_k)$ adjusts with changes in households' income and preferences for any of the K attributes.[38]

---

[37] Cropper et al. (1988) use simulation to illustrate how this allocation process would work, for a market with household exhibiting different tastes and income, using a maximum bid rule.

[38] Estimates from nonmarket valuation are used to make statements about the impacts of environmental and urban amenities on public welfare. In the particular case of the hedonic price function, however, Rosen (1974) notes that estimates from these models reflect only the capitalization rate of amenities in question, and do not necessarily reflect consumers' marginal willingness to pay (MWTP), except under a series of stringent and arguably inapplicable assumptions (for example, that the change in the dosage of the amenity is marginal in magnitude). Estimating MWTP from a hedonic price model thus requires structurally modeling a second step, which regresses MWTP for amenities and housing attributes on extra demand-side attributes. Other property valuation models have also been proposed as an alternative to the hedonic price model, including random utility, random bidding, and locational equilibrium models. For a more detailed discussion, see Palmquist (2005) and Shephard (1999). In general, however, assessments limited to estimating only the hedonic price model remain the most standard practice, for several reasons. First, non-marginal impacts can be estimated without a second stage, and interpreted as windfall gains or losses to property owners, if the researcher assumes that the attribute in question is sufficiently local, as to not shift the hedonic price function of the market as a whole (Palmquist, 2005). Second, the data requirements for welfare estimation are vast and often infeasible, as they require information from multiple housing markets, or demand-side data on the consumers of such housing. This is generally not accessible to researchers, whose datasets are confined to local assessment offices, multiple listing services, or other proprietary data providers. When household level data on consumers are available, such as the decennial Census, it is used at the expense of detailed geographic information, namely which households are most closely affected by amenities, as well as better information about the underlying value of the homes considered; in the case of the decennial Census, for example, provided home values are estimated by the respondents themselves. Second, welfare assessments require imposing additional assumptions about the housing market(s), the applicability of which is debatable. For example, modeling multiple housing markets requires assuming a common underlying preference for household attributes and amenities, which likely does not hold. The decision of households to locate in one market over another, for example, may be a result of endogenous sorting. In addition, assumptions

4.2.2 Omitted variable bias in hedonic price models

Omitted variables are a widely recognized problem, in the hedonic price literature. Their biasing effect on coefficients of interest are examined directly by Cropper et al. (1988, 1993), and Kuminoff et al. (2010), who use simulations to demonstrate the magnitude of bias generated from a misspecified model, depending on how the model is structured (e.g. linear, log-log, box-cox quadratic, etc.).[39]

Researchers seeking to mitigate omitted variables use more readily available data to control for more attributes, and more precisely. They also employ a wider variety of econometric techniques, such as spatial autocorrelative terms. The idea behind spatial autocorrelation is similar to autocorrelation with time series data, where the error terms for observations of consecutive time periods are correlated, in that properties within a vicinity of one another share spatially unobserved attributes which are reflected in the correlation of their error terms. Spatial models are applied to hedonic property models examining, among other things, open space (Irwin, 2002; Irwin and Bockstael, 2001; Acharya and Bennett, 2001; Geoghegan et al, 2003; Ready and Abdalla, 2005), air quality (Kim, Phipps, and Anselin, 2003; Anselin and Le Gallo, 2006; Anselin and

---

must be imposed regarding the structure of their underlying utility functions. These concerns apply also to other above mentioned alternatives to Rosen's (1974) structural model, including random utility and locational equilibrium models. Furthermore, models which consider location decisions are usually structured in a way as to ignore the dynamic element of decision-making in these markets; consumers' decisions are not all made simultaneously, and their choice sets at any given point in time comprise only the homes for sale during that period.

[39] Most research, to date, has concentrated on the standard hedonic price model rather than more structural alternatives, arguably for the reason that the ability to make welfare statements on the impact of amenities is of little value if estimates are obtained using a misspecified model. In other words, in lieu of the problems with structural models discussed in the previous footnote, much of the literature has to date concentrated on more accurately estimating the value of factors comprising the standard hedonic price model, by more satisfactorily addressing omitted variable bias.

Lozano-Gracia, 2008), noise externalities (Day et al, 2007), oil and gas facilities (Boxall et al, 2005), hazardous sites (Gawande and Jenkins-Smith, 2005; Brasington and Hite, 2005), wildfires (Mueller and Loomis, 2008), flood risk (Bin et al., 2008), and water quality (Leggett and Bockstael, 2000). In general, however, spatial models are not commonly exercised in the hedonic price literature. One reason is that they are often regarded as unnecessary, when the researcher has other geographic data about the surrounding area (e.g., the presence of surrounding structures, streams, and fields). Geographic information is also commonly used to combine parcel data with other datasets containing information on such factors as weather and pollution from local monitors, soil and stream quality, and school district quality.

Another alternative to spatial lags, for addressing omitted variables, is spatial fixed effects. These too, however, have shortcomings which render their use questionable in certain contexts. For one thing, there is the question as to how large the regions they encompass should be. If the areas are too broad, fixed effects will fail to control for more of the unobserved heterogeneity of smaller areas comprising them. If they are too small, one risks diluting the magnitude of the treatment effect itself, by incorporating it into the fixed effects encompassing homes sharing the treatment (for more on this topic, see Abbott and Klaiber, 2010). Another concern is that boundaries such as those of Census tracts may not identify natural neighborhoods. Identifying smaller entities such as block groups helps to correct for this. However, they may not be entirely successful at doing so.

Another method for addressing omitted variable bias is to estimate the hedonic price model using repeat sales, i.e. homes which have sold more than once, during the time window considered in the data. In this model, the difference in the prices between two sales of each property is regressed on any changes observed between the two sales. This purges the model of unobserved, potentially confounding variables which are time invariant. Repeat sales are highly regarded for their use in assessments where a home's exposure to an amenity measurably changes between two sales, in the absence of other confounding changes, by allowing one to compare the relative price appreciation between properties whose exposure changed by differing amounts.

Using repeat sales data also has limitations, however, which are well recognized in the literature. First, it assumes that covariates, namely housing attributes, do not change with time. This can be problematic if homes are remodeled or structurally altered between sales, unless the data allows one to control for when and how this occurs, or when there are unobserved changes in the neighborhood. Second, the samples of repeat sales available for estimation is often small for any given area. Third, and even more important, there are concerns that homes that sell multiple times over the study period (the only ones that are retained for the purposes of a repeat sale study) might be systematically and unobservably different from their single-sale counterparts, in which case the same is not representative of the full stock of homes. For these reasons, repeat sales tend not to be the sole identification strategy employed for an assessment, but used as a supplemental measure for gauging a result's robustness.

None of the three preceding methods discussed – spatial models, spatial fixed effects, and repeat sales – however, sufficiently address the potential concern of omitted, time-variant variables about neighborhood attributes which could covary with amenity changes. The value of a home which is sold near a recently built rail transit station, for example, reflects not only by the (dis)amenities provided by the station itself, but also the changes undergone by the surrounding area, as a result of its presence. The composition of the neighborhood residents, as well as the economic character of the neighborhood itself, may also change in ways not quite visible to the researcher.[40] It is very important, for this reason, that a researcher's assessment be structured in such a way that an amenity's treatment effect can be cleanly and credibly identified. This is the case with natural or quasi experiments such as those earlier noted in this chapter's introduction, whereby an amenity's introduction to or subtraction from a neighborhood can be plausibly regarded as a discrete, exogenous "shock".

4.2.3 Methamphetamine labs

Methamphetamines refer to a highly addictive stimulant which takes the form of a white, bitter, odorless, bitter-tasting powder which easily dissolves in water or alcohol. By increasing the release of dopamine in the brains of users, they can trigger long-lasting sensations of pleasure, energy, and alertness. Side effects associated with its abuse include hallucination, paranoia, aggressive behavior, and mood disturbances (National Institute on Drug Abuse, 2013). There are additional problems associated with withdrawal, as well as long term health. What make methamphetamines particularly unique and popular in the US illicit drug market is their relative ease of production, which can be accomplished with common household items. Operations typically involve the use of cold medicine containing ephedrine or pseudoephedrine, which can be purchased at most pharmacies.

[40]For the case of rail transit stations, see Vinha (2005) as an example.

In recent years, as clandestine labs have become more widely discovered, many local jurisdictions have passed legislation restricting large scale purchases of this drug, by requiring pharmacists to sell cold medicine over the counter and to track the purchasing behavior of frequent customers. What has followed is a decrease in large scale production, and a dispersal of activity among more producers. Previous work has examined the effect of supply shocks on the price and quality of methamphetamines. Dobkin and Nicosia (2009) consider the case of a 1995 intervention by the DEA which limited its availability to wholesale suppliers of clandestine labs, and find subsequent increases in prices and decreases in consumption, drug arrests, and the purity of the drug to be short lasting. The cause of these rebounds is unclear, although they surmise that supply disruptions were potentially offset by an increase in small-scale, home laboratory operations. In addition, there is evidence that methamphetamine use has expanded from the west towards the east, since the 1990s, particularly in the western, southern, and midwestern states wherein methamphetamine hospital related admissions rose by over 400% between 1992 and 2002 (Substance Abuse and Mental Health Services Administration, 2005). Some county officials have reported that methamphetamines are one of their biggest drug problems (National Association of Counties, 2005).

Production among residential clandestine labs is often small-scale, with most of the produced being personally consumed, and just enough left to sell in order to purchase additional inputs (Scott, 2002).[41] Even in such cases, however, methamphetamine labs pose public health risks to those residing in hosting premises, and possibly neighboring homes within their immediate surroundings. Meth production is a dangerous chemical

[41] With a small scale lab, only a few ounces of methamphetamine are typically made in a production cycle. This contrasts with larger labs, where production of at least 10 lbs per cycle is more commonly the case.

process, and can cause burns, fire, fumes, and/or explosions. It also creates environmentally toxic byproducts – at least five times the amount, in terms of weight, relative to the amount of drug produced (Bartos, 2005).

Following the seizure of a clandestine lab, the government takes responsibility for managing any potential threats to the immediate area, by handling the contaminants and byproducts which pose a threat to the public. This can involve evacuation of the surrounding area, for at least a few hours. Non-flammable contaminants within the inside of a home, however, such as traces of toxic chemicals, generally remain the owner's responsibility. While there remains much uncertainty in the medical community over the health effects of short and long term exposure to these types of chemicals, problems associated with the latter include cancer; damage to the brain, liver, and kidneys; and reproductive problems (US Environmental Protection Agency, 2011). Following its discovery, a lab is supposed to be documented in a federal register, which publicly discloses the address and date of such. Some states, counties, and municipalities also keep their own registers. Ideally, there should be complete overlap, so that a lab discovered in a city will be documented in the county, state, and federal registers, in addition to its own (assuming the state, and county have one). This turns out to not always be the case, though.

In at least 20 states, the interior of a home which hosted a former meth lab are required by mandate to undergo contamination cleanup before it can be sold or rented (Dewan and Brown, 2009), as well as removed from a register. This is a difficult and costly venture to undergo, however, and given the difficulty of proving that a home has been completely decontaminated, as well as the lack of national standards governing what makes for a sufficient cleanup, there are liability concerns involved the cleanup process. For these reasons, relatively few homes, once listed as a former clandestine lab, are removed from a register.

The presence of a meth lab in a neighborhood can be sometimes detected from a pungent odor, similar to ammonia or acetone (Benson, 2008). Additional signs of their presence include large amounts of trash or evidence of dumping in a yard or elsewhere nearby. Otherwise, however, it is often difficult to tell that a home is used as a meth lab, since equipment and materials are easily hidden and out of sight. Surprisingly, anecdotal evidence suggests that meth labs are typically discovered by government officials for reasons unrelated to neighbor suspicion, including domestic disturbances (see, e.g., Hrenda, 2016). Often, a lab is not discovered until *after* its hosts have vacated the premises, and other residents have moved in (Easter, 2010; Martin, 2014). The extent to which neighbors are typically aware of a meth lab's presence before its discovery by law enforcement remains an open question.

4.3 The model

I begin by considering a hedonic price model which takes the following form:

(4.3) $ln(P_{irt}) = \boldsymbol{X_{irt}}\beta_1 + \boldsymbol{Treated}_{irt}\beta_2 + \boldsymbol{Treated}*After_{irt}\beta_3 + \boldsymbol{Treated}*After*Pub_{irt}\beta_4 + \boldsymbol{\tau}_t + \boldsymbol{\gamma}_r + \varepsilon_{irt.}$

The dependent variable $ln(P_{irt})$ denotes the log price of property *i* sold at time *t* in region *r*. Regression covariates, denoted by the vector $\boldsymbol{X}_{irt}$, include characteristics of the home such as the log square acreage, log total living area, age of the structure, number of rooms and bathrooms, the style of the structure, etc.[42] I control for time-specific trends in real estate prices (e.g., the rise and subsequent fall of home prices during the 2000s) using month-year fixed effects, $\boldsymbol{\tau}_t$. Local fixed effects are controlled for at the Census block group level, as denoted by $\boldsymbol{\gamma}_r$.

As with Congdon-Hohman (2013), I use a difference-in-difference approach. The term $\boldsymbol{Treated}_{irt}$ is a vector of dummy variables denoting a home's proximity to several distance intervals of a discovered meth lab. It controls for unobserved, time invariant heterogeneity which differentiates properties located near discovered meth labs from further away properties. As with the block group fixed effects, β_2 mitigates concern about treated homes, for example, possessing time invariant qualities pertaining to quality and susceptibility to crime.

The average treatment effect on the treated is the coefficient on the interaction between ***Treated*** and *After*. My model assumes that, conditional on the neighborhood fixed effects (i.e., the block group fixed effects) and house characteristics, the treatment is as good as randomly assigned.

I wish to remind the reader that, consistent with a difference-in-difference design, my sample is comprised of home sale prices for homes that are regarded as treated and homes that are regarded as control. I assign homes to one or the other group on the basis of their distance from a discovery site. Specifically, as will become clear in the next section of this chapter, attention is restricted to homes within 400m, of one discovery during the time study period. (In Appendix 4, I consider specifications where the sample

[42] A full listing of covariates used in the model can be found in Table A.3 of Appendix 3.

is expanded to include homes near more than one discovery, and treatment is defined as the number of discoveries made within a defined distance interval. Such results report discoveries having a small, negative effect, which in some cases declines with distance from the discoveries themselves.)

It is possibly naive, however, to simply assume that buyers will be automatically aware of all previous nearby discoveries, when looking in a new neighborhood. Anecdotes suggest buyers sometimes have a difficult time determining even if the house that they are looking to buy previously hosted a meth lab. In some areas, sellers are legally required to disclose any knowledge of this. However, no law exists anywhere requiring the disclosure of nearby homes. Buyers do have the option, however, to inform themselves on the matter by researching homes in their neighborhood using meth lab registers.[43] One register which is maintained nationally, for example, is the National Clandestine Laboratory Register (NCLR), by the US Drug Enforcement Agency (DEA), which is populated by submissions from local agencies. In the case of Summit County, additional registers are maintained by the Summit County Sheriff (SCS), as well as the Akron Police Department (APD). However, this information is not necessarily accessible to homebuyers for periods preceding the publication of these registers.[44]

By comparing the three registers, I am able to identify discrepancies, as not every discovery appears on each applicable register. There are many lab discoveries in the SCS

[43] Similar to sex offender registers, there looks to be growing awareness of meth lab registers among homebuyers. See, for example, Nelson (2013).
[44] The DEA register was first published in December, 2006. The SCS and APD registers followed in May and September of 2008, respectively.

register, for example, which are not in the DEA or APD register.[45] By examining the registers individually, and determining when each register was made public, I am able to identify on which registers each discovery is listed and, based on the date the oldest register was made public, determine if information about the discovery was available to buyers when a property was purchased. In this chapter, I consider the role that this public information may have on the treatment effect by considering an additional coefficient, β_4, which considers the marginal treatment effect applicable to homes sold during periods when at least one of the register(s) disclosing the discovery's location was made available.

In this chapter, I also use estimate the model with repeat sales, which apply to properties that sell more than once over the study period. Assuming that covariates for property *i* remain do not vary between sales, equation (4.3) can be estimated with the following alternative:

(4.4) $\triangle ln(P_i) = \triangle \boldsymbol{Treated}{*}After_i\beta_3 + \triangle \boldsymbol{Treated}{*}After{*}Pub_i\, \beta_4 + \triangle \boldsymbol{\tau}_t + \triangle \varepsilon_i.$

This equation estimates the coefficients of interest, β_3 and β_4, by comparing the change in the values of properties which sold both before and after a meth lab's discovery between homes closest to discoveries to those farther away. This exploits the structure enabled by a panel dataset with a fixed effects specification, by reducing concerns about unobserved heterogeneity biasing the treatment effect.

[45] Up to the year 2012, the time span covered by each of the registers is as follows: for SCS, 2001-2012; for APD, 2002-2008; for DEA, 2000-2012.

4.4 Data

The data for this exercise originates from several sources. Information on property characteristics, transactions, and geographic location was provided by the Summit County Fiscal Office. I use registers provided by the DEA, SCS, and APD to identify the dates and locations of discovered meth labs. This allows me to map the location of these labs relative to properties in the county, using geographic information systems (GIS) software. The registers sometimes list more than one discovery for a defined address; in such cases, I consider only the earliest discovery. Data for Census tracts originates from GIS data from the US Census Bureau.

All observations, including those for repeat sales, are restricted to arms-length transactions, which are denoted in the data by a given code.[46] Attention is restricted to single family dwellings, as denoted by both the land use and style codes for the dwelling, [47] as well as homes with 10 or fewer bedrooms, and 5 or fewer bathrooms. Repeat sale observations are further filtered to omit those which 1) do not sell twice between 1 and 10 years of one another, and 2) to mitigate concerns of "flipped" homes and other outliers, exceed a 25 percent rate of annual appreciation. Finally, I trim both the single-sale and repeat-sale samples to eliminate outliers, dropping homes in the top and bottom one percent of the distribution of sales prices (single sale) and differences in sale prices (repeat sales).

[46] According to a systems analyst at the Summit County fiscal Office, arms-length transactions are denoted in the sales record file, used for this analysis, with a specific "sale validity code."
[47] Observations with inconsistencies between these two variables (e.g. single family according to the land use code, and condo according to the style code) were dropped.

Table 4.1 presents summary statistics, which compare the means of properties in my dataset, sold between 1999 and 2012, between those which were not located within 400m of any discovery, and those which were within 400m of at least one discovery. This second group consists only of sales which occurred prior to any of these discoveries. Table 4.1 presents the first piece of evidence that a discovery is not a purely exogenous phenomenon, as these two groups of homes are quite from one another, on the basis of observable attributes. Those over 400m of any discovery have much larger sale values, larger amounts of acreage and living area, and more bedrooms and bathrooms than homes located closer to discoveries. Homes in the former group are also about half the age of homes in the latter group, on average. These findings are consistent with Congdon-Hohman (2013), who similarly finds with his dataset that observable characteristics of homes between these two groups are different.

Table 4.1. differences in means between properties within or over 400m of at least one discovery

Variable	no discoveries within 400m Mean	1+ discoveries within 400m Mean	t-stat
Sale price	195282.990	86063.950	-174.865
Stories	1.518	1.461	-13.827
Total acreage	0.554	0.220	-74.188
Total living area (sq. ft)	1988.542	1268.159	-146.438
Total bedrooms	3.300	2.841	-75.847
Total bathrooms	1.755	1.181	-131.667
Total half bathrooms	0.659	0.263	-95.364
Age	30.927	60.256	133.498
Sold in same year built	0.098	0.021	-45.446
Exterior wall: frame	0.114	0.157	14.128
Exterior wall: masonry and frame	0.005	0.004	-2.901
Exterior wall: aluminum/vinyl	0.770	0.757	-3.394
Style: colonial	0.468	0.414	-12.608
Style: ranch	0.274	0.283	2.356
Style: cape cod	0.111	0.164	17.504

In Table 4.2, I further explore what characteristics are associated with neighborhoods with meth discoveries, using data on block groups in Summit County from the Census in the year 2000, which precedes the year of the first discovery in my data, 2001. I regress the number of discoveries counted for each block group on mean and median characteristics of that group. These results indicates that more discoveries occur in block groups with lower median time spent in a residence and lower median home values. More numerous discoveries are also positively correlated with the percentage of the population that is less educated, white, living in renter-occupied housing, and living in a three to four person housing unit. As with the previous table, Table 4.2 supports the notion that discoveries are not a purely random phenomenon.

Table 4.2. Predicting the number of meth discoveries per Census block group

VARIABLES	
log median household income	0.630
	(0.625)
median home age	0.00869
	(0.00807)
median time in home	-0.0660**
	(0.0286)
log median gross rent	0.101
	(0.310)
log median home value	-1.598***
	(0.471)
% in labor force	-0.00235
	(0.0129)
% unemployed	-0.0301
	(0.0256)
% educ: below HS	0.0641***
	(0.0242)
% educ: HS	-0.0265
	(0.0165)
% educ: associates	-0.0137
	(0.0342)
% educ: bachelors	-0.0627**
	(0.0271)
% population: white, nonhispanic	0.0496***
	(0.00609)
% population: in renter occupied homes	0.0361***
	(0.0138)
% population: in family households	0.0185
	(0.0125)
% population: in non-family households	-0.0280

	(0.0227)
% vacant housing units	-0.00485
	(0.0337)
% households: 3-4 person	0.0282*
	(0.0163)
% households: 5+ person	-0.00798
	(0.0341)
Constant	8.035
	(7.381)
Observations	460
R-squared	0.374

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The sample omits 17 block groups, due to censored data for variables such as gross rent.

Congdon-Hohman (2013) argues that, while meth discoveries are not a purely exogenous phenomenon, they are locally exogenous in that while some neighborhoods are more likely to have them, the precise location within the neighborhood itself is not easy to predict. In his assessment, he thus includes in his sample only homes with only one discovery within 400m (one quarter mile) during his sample period (January, 2002 to March, 2009), and uses a difference-in-difference approach to argue that the treatment effect can be estimated by comparing homes, within this group, which are located closer to that discovery with homes located further away. He does not consider sales which occurred near more than one discovery, arguing that it becomes more difficult to estimate the effect when there are additional discoveries to consider. He also reduces his sample to sales which occurred within one or two years of the neighborhood discovery. Following his rationale, I consider a similar approach by narrowing my sample to sales which occurred within 400m of only one discovery during my study period. My approach is different here, in that it includes information from two additional registers and a larger time period, changing the composition of the sample I identify as within 400m of one discovery.

In Table 4.3, I consider the same summary statistics presented in Table 4.1, using only homes I identify as within 400m of one discovery, before the discovery itself occurred. Comparing the means between these two groups, we see fewer statistically significant differences in the observable attributes of the homes. When such statistically significant differences do exist, the t-statistic is much lower, and the magnitude of the difference is much lower than the magnitudes observed in Table 4.1. Table 4.3 shows that the average sales price of homes located within 200m of a discovery are about $6000 lower than for those 200m to 400m of a discovery. Properties that are located farther from a discovery site also have on average a somewhat larger living area, and fewer half bathrooms.

These finding generally support the notion that the sample should be narrowed to homes located within discovery-prone neighborhoods, and that the “treatment” homes should be the closest to the discovery sites, whereas the control group comprises homes which are relatively farther from the discoveries in those very neighborhoods.

Table 4.3. differences in means between properties within 400m of one discovery

Variable	one discovery within 200m Mean	one discovery between 200m-400m Mean	t-stat
Sale price	103119.580	109831.250	-4.645
Stories	1.393	1.385	0.621
Total acreage	0.300	0.316	-1.332
Total living area (sq. ft)	1361.356	1388.037	-2.034
Total bedrooms	2.914	2.922	-0.438
Total bathrooms	1.273	1.283	-0.701
Total half bathrooms	0.345	0.383	-2.784
Age	52.422	51.661	1.109
Sold in same year built	0.023	0.024	-0.288
Exterior wall: frame	0.167	0.139	2.808
Exterior wall: masonry and frame	0.005	0.006	-0.522
Exterior wall: aluminum/vinyl	0.738	0.749	-0.941
Style: colonial	0.345	0.335	0.765
Style: ranch	0.304	0.327	-1.819
Style: cape cod	0.197	0.189	0.796

4.5 Results

I begin with by considering regression results in Table 4.4 using the hedonic price model specified in equation (4.3). For Table 4.4, I follow a procedure similar to Congdon-Hohman’s, by limiting the sample to properties within 400m of only one discovery during the sample period which are sold one year before or after a discovery – with the exception of homes sold within 30 calendar days after a discovery, which are excluded from the sample. [48] In the interest of space, only the coefficient estimates for β_3 and β_4 are displayed.[49] I begin in column (1) by defining treatment as the closest exposure to any site denoted in any of the three registers used for this analysis. For these results, I use two distance increments to define treatment: within 200m, and within 200m to 400m of a discovery. Column (1) considers the case for treatment associated with all discoveries listed in the three registers. The treatment effect for homes within the closest proximity, 200m, is positive, interestingly, although not statistically significant, suggesting that properties sold after a discovery do not incur a negative loss. For homes sold after a discovery within 200m to 400m instead, the finding is similar, indicating that these homes are not affected by a discovery, either.

Table 4.4. Hedonic price model estimates, for homes within 400m of only one discovery and sold within one year before or after the discovery

	(1)	(2)	(3)	(4)
	all meth discoveries		discoveries specifically denoted as "labs"	
Treated x After, 0-200m	0.0510	0.00174	0.0288	-0.00328
	(0.0418)	(0.0454)	(0.0404)	(0.0495)
Treated x After, 0-200m x Public Register		0.149*		0.0912
		(0.0775)		(0.0688)
Treated x After, 200-400m	0.0184	0.0284	0.0290	0.00805
	(0.0218)	(0.0263)	(0.0241)	(0.0314)
Treated x After, 200-400m x Public Register		-0.0424		0.0615
		(0.0478)		(0.0525)

[48] My sample selection criterion is a little more restrictive than Congdon-Hohman's, who considers homes sold within one year of one discovery. This latter approach potentially includes neighborhoods which potentially observe more than one discovery over a larger time window. Given the earlier observed differences between neighborhoods with zero versus multiple discoveries from Table 4.1, and concerns about potential lingering, stigmatizing effects that older discoveries still listed on a register could have on future property sales, I focus on neighborhoods observing their first, and only, discovery.

[49] Estimates for model covariates are included in Table A.3 of Appendix 3.

Observations	977	977	938	938
R-squared	0.876	0.878	0.886	0.887

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

In column (2), I refine the specification by including interaction terms which correspond to β_4 from equation (4.3), differentiating between sales sold before and after the public disclosure of the nearby site via a register. This interaction term can thus be interpreted as the additional treatment effect of proximity to a discovery when information about the discovery is more readily accessible to the homebuyer, at the time of purchase. Column (2)'s results do not suggest that this additional information produces a negative effect on property values. If anything, it suggests the opposite: homes within 200m of a discovery that were sold when at least one register was accessible to reference the discovery appear to appreciate by over 10%. For homes between 200m and 400m, the additive effect of the two coefficients, β_3 and β_4 is negative, but insignificant.

Part of the reason for column (1) and (2)'s results may relate to the small sample size used to estimate the treatment effects, with just fewer than 1,000 observations. Ideally, a hedonic price model relies on a large sample size to generate estimates from – at least several thousand observations. Part of it, however, may also pertain to the uncertainty concerning just what exactly constitutes a meth discovery. In columns (3) and (4), I subject my definition of a discovery to more scrutiny. One of the benefits of using the SCS register in this assessment is that most of its entries include brief commentary, offering a crude idea of the type of discovery that was made. Examples of words and phrases used include "meth lab[s]," "dump site[s]," discoveries of “chemical[s],” “box lab[s]”, “one pot[s]”, etc. Ideally, we prefer to differentiate between sites which typify what we associate with a discovered meth lab, and less relevant discoveries. Conversations with SCS personnel suggest there are many connotations associated with documented meth sites, concerning the magnitude and presence of operations. A “one pot,” for example, refers to a small scale, “shake and bake” type of operation, in which the meth ingredients are combined and mixed in a small container, typically a two liter bottle of soda. Thus, a “one pot” may refer to nothing more than the discovery of a two liter bottle of meth, or its byproducts, near the given address (e.g., in a dumpster or alongside the road). We might expect these types of documented discoveries to have a smaller impact on property values, as the stigma about neighborhood crime generated by their discovery is arguably lower. For example, it is hard to determine just from where a “one pot” originated, so it need not necessarily attest to the quality of the neighborhood itself. Furthermore, the small scale of the siting may be of a lesser public health concern to the surrounding neighborhood. Columns (3) and (4) thus narrows the definition of treated homes to those proximal to sites which the SCS, the largest of the three registers, suggests are actual labs.

Columns (3) and (4)'s results tell a similar story to columns (1) and (2). This time, however, in the case of both columns the treatment effect is insignificant for homes within closest proximity of a meth lab, in addition to the insignificance observed for homes within 200m to 400m of a discovery.

In Table 4.5, I alternatively consider specifications which include all properties within 400m of one discovery, irrespective of the time of sale. This serves two purposes: first, it increases the sample size. It is possible that the lack of significant findings in

Table 4.4 are attributable to the small number of observations, relative to what is desirable for a hedonic property model. Second, it allows for the treatment effect to persist for observations beyond one year. This is potentially justifiable, for the reason that discoveries could have a stigmatizing effect on neighborhoods which persists for a long time, since a discovery, once listed on the register, is rarely removed. Table 4.5's results, however, tell a similar story to that of Table 4.4. Again, no statistically significant change is observed among properties treated by a discovery, in either distance increment considered. This is the case both for when discoveries are defined to include all discoveries listed in the registers, and for when they are judged to be meth labs.[50]

Table 4.5. Hedonic price model estimates, for homes within 400m of only one discovery

	(1)	(2)	(3)	(4)
	all meth discoveries		discoveries specifically denoted as "labs"	
Treated x After, 0-200m	0.00544	-0.0118	0.0235	0.0112
	(0.0139)	(0.0183)	(0.0148)	(0.0212)
Treated x After, 0-200m x Public Register		0.0380		0.0261
		(0.0232)		(0.0260)
Treated x After, 200-400m	0.0132	0.00378	0.0147	0.00263
	(0.00924)	(0.0104)	(0.00996)	(0.0123)
Treated x After, 200-400m x Public Register		0.0265*		0.0269
		(0.0149)		(0.0166)
Observations	9,041	9,041	9,685	9,685
R-squared	0.749	0.749	0.735	0.735

[50] Note that, unlike Table 4.4, the sample size for homes near meth labs is larger than those for all discoveries. This is due to the fact that, because there are fewer meth labs than total discoveries, there are more properties located near more than one discovery than properties located near more than one meth lab.

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.6 produces results for the four previous specifications considered in Table 4.5, using repeat sales, as outlined in equation (4.2).[51] Before proceeding, it is worth noting the substantial decrease in the sample size. Within my data, only 1,780-1,887 properties have a second arms-length sale between 1999 and 2012, using the criteria previously defined in section 4.3 for inclusion in the sample.

For this reason, I take a similar approach to that of Table 4.5 and expand my sample to include all property sales within 400m of one discovery (whether or not sales occurred within a year of the discovery). Column (1) of Table 4.6 tells a different story from those of Tables 4.4 and 4.5: it suggests that properties farther away from meth discoveries are appreciating faster – this time, in the magnitude of about 5.6% - compared to the treated homes, the values of which do not appear to rise at all. It is possible that the discoveries may have simply prevented homes in closer proximity from appreciating as much as they might have, otherwise.

In Column (2), we also observe a different finding. When I allow for the treatment effect to vary for periods before and after registers were available to reference these discoveries at the time of sale, the results suggest a substantially negative treatment effect for homes within 200m of a discovery, in the magnitude of a little over 20%. For homes within 200m to 400m of a discovery, the treatment effect does not appear to be

[51] A common practice in the repeat sale literature as far back as Case and Shiller (1987) is to assume heteroscedastic error terms, which are a function of the length of time between the two sales. I account for potential heteroscedasticity, among my repeat sale specifications, by using robust standard errors.

that sensitive to the presence of registers. These findings are similar to those reported in columns (3) and (4), which if anything suggest an amplified effect when I consider only homes near discoveries identified as meth labs. It is difficult to know to what extent the magnitudes of the treatment effects generated in columns (2) and (4) of Table 4.6 – between 22% and 32% - are plausible.[52] As with Table 4.4, the small sample size may be a concern, since I am using less than 2,000 observations for any given specification. Given the questionable magnitude of the estimates from columns (2) and (4), as well as its inconsistency with my estimates from Tables 4.4 and 4.5, I do not regard such negative treatment effects to be a particularly robust finding.

Table 4.6. Hedonic price model estimates, using repeat sales, for homes within 400m of only one discovery

	(1)	(2)	(3)	(4)
	all meth discoveries		discoveries specifically denoted as "labs"	
Treated x After, 0-200m	0.0139	0.0219	0.0479*	0.0641**
	(0.0239)	(0.0241)	(0.0289)	(0.0296)
Treated x After, 0-200m x Public Register		-0.237**		-0.380***
		(0.115)		(0.0886)
Treated x After, 200-400m	0.0562***	0.0579***	0.0349*	0.0364*
	(0.0206)	(0.0207)	(0.0208)	(0.0209)
Treated x After, 200-400m x Public Register		-0.00311		-0.0270
		(0.0756)		(0.112)
Observations	1,780	1,780	1,887	1,887
R-squared	0.224	0.225	0.220	0.222

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

[52] These estimates of the treatment effect are higher than those of Congdon-Hohman (2013), which range from 10% to 19% for homes most proximal to a discover, and sold within one year of such.

4.6 Conclusions

This chapter considers an application of nonmarket valuation, by examining the effect of methamphetamine discoveries on property values. Following from previous literature, I revisit the study area of Summit County, Ohio using a more complete listing of discoveries from three registers covering the area than was previously considered, and a different set of selection criteria for properties considered in the model. Following with the previous literature, in recognizing that methamphetamine laboratories are not a purely exogenous phenomenon, I limit my assessment only to homes in neighborhoods affected by just one discovery.

I do not find strong evidence suggesting that home located near these discoveries observe negative depreciation in property values. A few results suggest a potentially negative effect, either directly or in terms of relatively higher properties for properties that are somewhat further away from such discoveries. In such cases when a negative effect is observed, it is for homes sold during periods that registers were available to reference these discoveries. However, it is unclear if these findings are credible, since the results are not robust to alternative specifications – namely, those using single sale observations – as well as the relatively small sample sizes available. Narrowing my list of methamphetamine discoveries to those which do not appear to be dump sites, using data from the Summit County register, serves to amplify the magnitude of the treatment effect, when such an effect is observed.

While my results are inconclusive, in revisiting this topic, this paper emphasizes several important factors to consider, when studying crimes of this nature, and conducting nonmarket valuation. First, treatment effects might be heterogeneous, depending on the type of shock observed by a neighborhood. In the case of the DEA register, used alone in the previous assessment, methamphetamine discoveries of all types – be they labs or dump sites – are listed in a way as to be indistinguishable from one another. In the case of the Summit County register, while the descriptions given for the listed discoveries are imperfect, they are useful for shedding light on this topic by offering a somewhat better sense of the type of discovery observed by a neighborhood. Second, information is important. If methamphetamine discoveries, and other negative shocks impose negative externalities on their neighborhoods through changes in property values, the question remains how homebuyers are made aware of their presence. Stories from local news outlets are one possibility; however, the effect of this information outlet is difficult to gauge. In the case of this particular negative externality, I consider the information effect of registers, by differentiating the treatment effect between periods before and after they were available. Given the inconsistent findings from this exercise, it is unclear if registers serve to make homebuyers more aware. Future research may be able to provide further insight on this, by gauging the awareness of homeowners in, and prospective buyers of, neighborhoods affected by methamphetamine discoveries. Third, sample size and area are important considerations for hedonic price models. In an assessment of the type considered, where the sample is confined to homes in the neighborhoods where only one shock is observed over the study period, results are likely to be sensitive to the inclusion and omission of different neighborhoods. As my assessment includes discoveries from two extra registers, relative to Congdon-Hohman (2013), this results in my counting additional meth labs in some neighborhoods where one was previously assumed to exist, as well as counting meth labs in some areas where none were previously accounted for.

Appendix

Appendix 1. Additional tables and figures for Chapter 2

Figure A.1. Distribution of RECS and AHS observations, using the state/state-group identifiers from RECS 2009

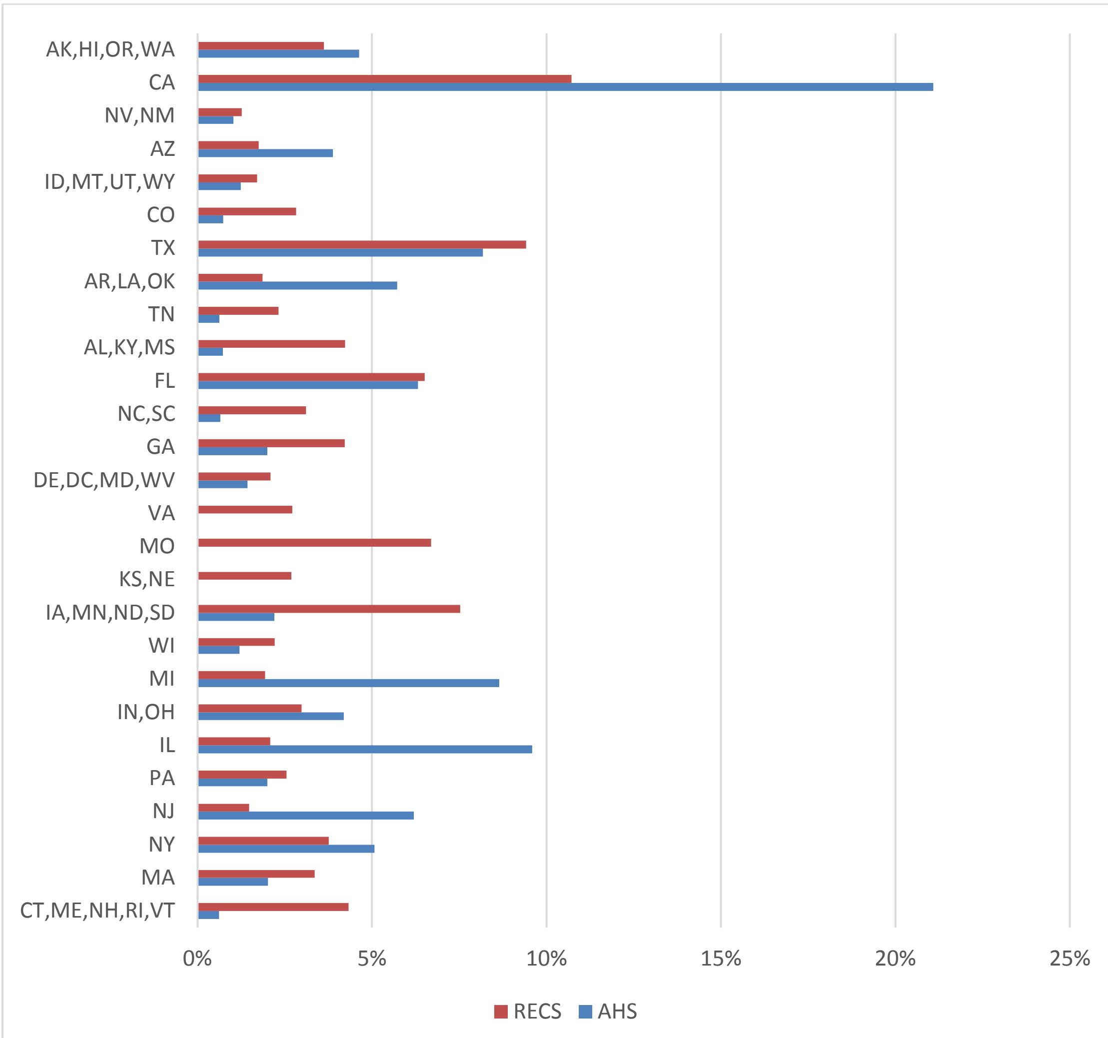


Table A.1. Summary statistics for the three data sources (RECS, AHS, Form 861); year 2009 only

Panel A. Number of cross sectional units

RECS	AHS	Form 861
6070	11180	2969

Panel B. Annual electricity consumption per household (kWh)

	mean	s.d.	min	max
RECS	12721.67	6882.009	2201	38805
AHS	12803.57	7174.527	1747.51	38656.82
Form 861	11496.44	3150.301	4819.355	21103.45

Panel C. Average price of electricity per kWh (2009 $)

	mean	s.d.	min	max
RECS	0.122334	0.034277	0.026188	0.291703
AHS	0.124944	0.031727	0.065446	0.235759
Form 861	0.103329	0.025764	0.026428	0.327753

Panel D. Average square footage of home

	mean	s.d.	min	max
RECS	2624.671	1330.242	409	9808
AHS	1924.21	1037.351	400	10000

Panel E. Ownership of capital stock

	AC		Electric heating	
	mean	s.d.	mean	s.d.
RECS	0.858484	0.348581	0.289292	0.453471
AHS	0.838819	0.367714	0.238372	0.426107

Panel F. Climate

	Heating degree days				Cooling degree days			
	mean	s.d.	min	max	mean	s.d.	min	max
RECS	4295.909	2303.401	6	11679	1392.588	1100.015	1	5407
AHS	3921.629	2288.992	140.175	8197.99	1410.956	1203.355	57.11429	5001.523

Panel G. Geographic distribution (%)

	AHS	RECS
CT,MA,ME,NH,RI,VT	1.79	7.69
NJ,NY,PA	14.49	7.79
IL,IN,MI,OH,WI	22.30	9.21
IA,KS,MN,MO,ND,NE,SD	1.62	16.90
DC,DE,FL,GA,MD,NC,SC,VA,WV	6.96	18.65
AL,KY,MS,TN	0.97	6.56
AR,LA,OK,TX	22.29	11.29
CO,ID,MT,UT,WY,AZ,NM,NV	4.83	7.56
AK,CA,HI,OR,WA	24.75	14.35

Appendix 2. Additional tables and figures for Chapter 3

Table A.2. Travel demand regression results, not corrected for self-selection

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log price (pence) per 100km	-0.344***	-0.489***	-0.335***	-0.330***	-0.388***	-0.472***	-0.427***
	(0.0553)	(0.0490)	(0.0823)	(0.0722)	(0.0928)	(0.132)	(0.106)
log price (pence) per 100km x (130-149g/km emissions)					0.00109	0.00339	0.00848*
					(0.00447)	(0.00876)	(0.00510)
log price (pence) per 100km x (150-199g/km emissions)					0.00568	0.0142	0.0126*
					(0.00591)	(0.0102)	(0.00662)
log price (pence) per 100km x (200-249g/km emissions)					0.00737	0.0192	0.0160
					(0.00893)	(0.0139)	(0.0103)
log price (pence) per 100km x (>=250g/km emissions)					-0.00197	0.0137	0.00494
					(0.0119)	(0.0178)	(0.0143)
rural	0.125***	0.123***	0.135***	0.112***	0.125***	0.135***	0.112***
	(0.0130)	(0.0130)	(0.0188)	(0.0181)	(0.0130)	(0.0188)	(0.0181)
diesel	0.196***	0.139***	0.204***	0.194***	0.198***	0.205***	0.199***
	(0.0208)	(0.0188)	(0.0337)	(0.0261)	(0.0214)	(0.0348)	(0.0265)
log engine capacity	0.384***	0.483***	0.432***	0.324***	0.383***	0.422***	0.317***
	(0.0472)	(0.0441)	(0.0679)	(0.0657)	(0.0475)	(0.0682)	(0.0672)
age of car 1-2 years	0.0489***	0.0460***	0.0396**	0.0554***	0.0485***	0.0409**	0.0536***
	(0.0116)	(0.0114)	(0.0163)	(0.0160)	(0.0116)	(0.0163)	(0.0160)
small/medium car	0.0940***	0.0941***	0.0987***	0.0850***	0.0910***	0.0930***	0.0771***
	(0.0190)	(0.0190)	(0.0263)	(0.0278)	(0.0191)	(0.0265)	(0.0283)
medium car	0.239***	0.232***	0.249***	0.227***	0.232***	0.238***	0.214***
	(0.0216)	(0.0216)	(0.0307)	(0.0304)	(0.0220)	(0.0317)	(0.0312)
large car	0.335***	0.340***	0.334***	0.328***	0.326***	0.321***	0.311***

	(0.0265)	(0.0265)	(0.0375)	(0.0374)	(0.0267)	(0.0379)	(0.0380)
land rover / jeep	0.146***	0.170***	0.156***	0.135***	0.149***	0.155***	0.137***
	(0.0357)	(0.0355)	(0.0511)	(0.0494)	(0.0357)	(0.0515)	(0.0493)
household income in 2nd quintile	-0.00937	-0.0106	-0.0346	0.0152	-0.00804	-0.0293	0.0158
	(0.0271)	(0.0271)	(0.0387)	(0.0380)	(0.0271)	(0.0387)	(0.0381)
household income in 3rd quintile	0.0754***	0.0726***	0.0835**	0.0678*	0.0764***	0.0864**	0.0686*
	(0.0256)	(0.0257)	(0.0370)	(0.0357)	(0.0256)	(0.0370)	(0.0358)
household income in 4th quintile	0.144***	0.142***	0.171***	0.118***	0.144***	0.172***	0.119***
	(0.0252)	(0.0253)	(0.0359)	(0.0358)	(0.0252)	(0.0359)	(0.0359)
household income in 5th quintile	0.186***	0.185***	0.215***	0.160***	0.187***	0.218***	0.161***
	(0.0254)	(0.0255)	(0.0367)	(0.0357)	(0.0254)	(0.0367)	(0.0358)
Family household, with adult child(ren)	0.0241	0.0191	0.0304	0.0152	0.0243	0.0309	0.0168
	(0.0210)	(0.0211)	(0.0313)	(0.0287)	(0.0210)	(0.0313)	(0.0287)
Pensioner household	-0.201***	-0.212***	-0.192***	-0.212***	-0.201***	-0.192***	-0.211***
	(0.0247)	(0.0247)	(0.0372)	(0.0330)	(0.0247)	(0.0371)	(0.0330)
Other household	0.00385	-0.000204	0.0232	-0.0161	0.00379	0.0240	-0.0166
	(0.0145)	(0.0145)	(0.0204)	(0.0207)	(0.0145)	(0.0203)	(0.0207)
NW & Merseyside	-0.00868	-0.00808	-0.0657	0.0433	-0.0103	-0.0681	0.0422
	(0.0287)	(0.0288)	(0.0423)	(0.0393)	(0.0287)	(0.0423)	(0.0392)
Yorkshire & Humberside	0.00485	0.00630	-0.0801*	0.0826**	0.00340	-0.0828*	0.0807*
	(0.0302)	(0.0302)	(0.0443)	(0.0415)	(0.0302)	(0.0443)	(0.0414)
East Midlands	0.0109	0.0132	0.0138	0.00958	0.00915	0.00906	0.00951
	(0.0308)	(0.0309)	(0.0445)	(0.0426)	(0.0308)	(0.0446)	(0.0425)
West Midlands	-0.00653	-0.00366	-0.0530	0.0380	-0.00778	-0.0563	0.0384
	(0.0296)	(0.0297)	(0.0422)	(0.0417)	(0.0296)	(0.0423)	(0.0416)
Eastern Region	-0.0203	-0.0180	-0.0764*	0.0315	-0.0216	-0.0805*	0.0330
	(0.0295)	(0.0296)	(0.0432)	(0.0405)	(0.0295)	(0.0433)	(0.0405)

Greater London	-0.249***	-0.242***	-0.271***	-0.238***	-0.249***	-0.273***	-0.237***
	(0.0356)	(0.0357)	(0.0506)	(0.0504)	(0.0356)	(0.0507)	(0.0504)
South East Region	-0.0349	-0.0329	-0.0539	-0.0186	-0.0366	-0.0586	-0.0178
	(0.0281)	(0.0282)	(0.0412)	(0.0386)	(0.0281)	(0.0414)	(0.0385)
South West Region	-0.00485	0.000416	-0.0595	0.0471	-0.00576	-0.0606	0.0466
	(0.0307)	(0.0307)	(0.0445)	(0.0425)	(0.0307)	(0.0445)	(0.0425)
Wales	0.0367	0.0372	-0.0312	0.0945**	0.0354	-0.0322	0.0945**
	(0.0345)	(0.0345)	(0.0504)	(0.0474)	(0.0344)	(0.0504)	(0.0474)
Scotland	-0.0394	-0.0415	-0.0749*	-0.0101	-0.0408	-0.0807*	-0.00991
	(0.0296)	(0.0297)	(0.0436)	(0.0405)	(0.0295)	(0.0436)	(0.0404)
Time from bus to rail station >14 mins.	-0.0290**	-0.0322**	-0.0513***	-0.0101	-0.0290**	-0.0511**	-0.00976
	(0.0139)	(0.0139)	(0.0199)	(0.0195)	(0.0139)	(0.0198)	(0.0195)
Walk time to rail station >26 mins	0.0273*	0.0303**	0.0315	0.0233	0.0272*	0.0309	0.0237
	(0.0148)	(0.0148)	(0.0212)	(0.0207)	(0.0148)	(0.0212)	(0.0207)
Walk time to bus stop >6 mins.	-0.00944	-0.00943	-0.00939	-0.0139	-0.00863	-0.00739	-0.0132
	(0.0155)	(0.0155)	(0.0221)	(0.0217)	(0.0154)	(0.0221)	(0.0216)
Household owns home	-0.0257	-0.0200	-0.0281	-0.0205	-0.0252	-0.0279	-0.0208
	(0.0214)	(0.0213)	(0.0332)	(0.0280)	(0.0214)	(0.0332)	(0.0280)
Socioeconomic group: clerical	-0.0442***	-0.0464***	-0.0388*	-0.0505**	-0.0442***	-0.0385*	-0.0507**
	(0.0152)	(0.0152)	(0.0216)	(0.0213)	(0.0152)	(0.0216)	(0.0213)
Socioeconomic group: skilled manual	-0.144***	-0.141***	-0.109***	-0.181***	-0.143***	-0.109***	-0.181***
	(0.0198)	(0.0198)	(0.0268)	(0.0291)	(0.0198)	(0.0268)	(0.0290)
Other socioeconomic group / other manual	-0.131***	-0.130***	-0.0876***	-0.182***	-0.131***	-0.0883***	-0.183***
	(0.0245)	(0.0247)	(0.0338)	(0.0357)	(0.0245)	(0.0338)	(0.0355)
Socioeconomic group: retired	-0.275***	-0.271***	-0.262***	-0.287***	-0.275***	-0.262***	-0.289***
	(0.0233)	(0.0233)	(0.0355)	(0.0308)	(0.0233)	(0.0354)	(0.0308)
Socioeconomic group: other economically inactive	-0.231***	-0.230***	-0.162**	-0.298***	-0.230***	-0.158**	-0.299***
	(0.0485)	(0.0486)	(0.0683)	(0.0692)	(0.0486)	(0.0685)	(0.0696)
2 persons in household	-0.0187	-0.0210	-0.0259	-0.0133	-0.0193	-0.0265	-0.0147

with driver's license	(0.0145)	(0.0146)	(0.0204)	(0.0207)	(0.0145)	(0.0204)	(0.0207)
3 persons in household	0.0156	0.0103	0.0133	0.0204	0.0150	0.0130	0.0172
with driver's license	(0.0262)	(0.0263)	(0.0390)	(0.0357)	(0.0262)	(0.0390)	(0.0356)
4+ persons in household	0.0176	0.00895	0.0104	0.0295	0.0182	0.0137	0.0265
with driver's license	(0.0359)	(0.0359)	(0.0551)	(0.0479)	(0.0359)	(0.0554)	(0.0479)
Survey year: 2002	0.101*				0.101*		
	(0.0579)				(0.0579)		
Survey year: 2003	0.0186				0.0181		
	(0.0470)				(0.0471)		
Survey year: 2004	0.0332				0.0361		
	(0.0497)				(0.0496)		
Survey year: 2005	0.00913				0.0143		
	(0.0506)				(0.0508)		
Survey year: 2006	-0.0295				-0.0222		
	(0.0502)				(0.0507)		
Survey year: 2007	-0.00281				0.00357		
	(0.0505)				(0.0509)		
Survey year: 2008	-0.0510				-0.0374		
	(0.0519)				(0.0546)		
Survey year: 2009	-0.0784				-0.0719		
	(0.0525)				(0.0528)		
Survey year: 2010	-0.0980*				-0.0810		
	(0.0593)				(0.0628)		
Constant	8.220***	8.428***	7.858***	8.537***	8.474***	8.731***	9.148***
	(0.277)	(0.270)	(0.389)	(0.373)	(0.544)	(0.761)	(0.662)
Observations	10,689	10,689	5,050	5,639	10,689	5,050	5,639
R-squared	0.276	0.271	0.290	0.264	0.276	0.291	0.265
Sample period	2002-2010	2002-2010	2002-2005	2006-2010	2002-2010	2002-2005	2006-2010

Standard errors in parentheses, and are clustered at the household level; *** p<0.01, ** p<0.05, *

$p<0.1$

Appendix 3. Additional tables for Chapter 4

Appendix Table 3. Additional regression results for baseline model from column (1) of Table 4.5

Treated, 0m-200m	-0.0181**
	(0.00821)
Stories	0.0123
	(0.0227)
Log total acreage	0.0653***
	(0.00901)
Log total living area	0.406***
	(0.0181)
Total bedrooms	0.0246***
	(0.00548)
Total bathrooms	0.0569***
	(0.00744)
Total half bathrooms	0.0623***
	(0.00633)
Age	-0.00683***
	(0.000504)
Age Squared	2.39e-05***
	(4.97e-06)
Sold in same year built	-0.502***
	(0.0534)
Exterior wall: brick	0.0546***
	(0.0116)
Exterior wall: masonry and frame	0.0172
	(0.0506)
Exterior wall: block	-0.188**
	(0.0740)
Exterior wall: stucco	0.0671
	(0.0423)
Exterior wall: aluminum/vinyl	0.00975
	(0.00847)
Exterior wall: stone	-0.0858
	(0.0542)
Exterior wall: masonry and vinyl	0.0434
	(0.0325)

Style code: multilevel	-0.0705
	(0.0985)
Style code: bi-level	-0.0204
	(0.0413)
Style code: log home	0.205***
	(0.0646)
Style code: modular	-0.136
	(0.0868)
Style code: colonial	0.0182
	(0.0342)
Style code: ranch	0.00920
	(0.0390)
Style code: bungalow	-0.00445
	(0.0388)
Style code: raised ranch	0.00615
	(0.0488)
Style code: contemporary	0.104**
	(0.0508)
Style code: tudor	0.141**
	(0.0616)
Style code: split level	-0.00696
	(0.0406)
Style code: cape cod	0.0108
	(0.0377)
Feb. 1999	-0.0523
	(0.0676)
Mar. 1999	0.0119
	(0.0573)
Apr. 1999	0.00142
	(0.0578)
May-99	0.0531
	(0.0539)
Jun. 1999	0.0425
	(0.0550)
Jul. 1999	0.0733
	(0.0542)
Aug. 1999	0.125**
	(0.0546)
Sep. 1999	0.0515
	(0.0542)

Oct. 1999	0.0983*
	(0.0545)
Nov. 1999	0.0641
	(0.0573)
Dec. 1999	0.164***
	(0.0523)
Jan. 2000	0.145**
	(0.0677)
Feb. 2000	0.0395
	(0.0647)
Mar. 2000	0.111**
	(0.0530)
Apr. 2000	0.133***
	(0.0515)
May-00	0.182***
	(0.0493)
Jun. 2000	0.158***
	(0.0508)
Jul. 2000	0.0677
	(0.0602)
Aug. 2000	0.165***
	(0.0522)
Sep. 2000	0.177***
	(0.0526)
Oct. 2000	0.147***
	(0.0566)
Nov. 2000	0.138**
	(0.0553)
Dec. 2000	0.166***
	(0.0599)
Jan. 2001	0.0649
	(0.0617)
Feb. 2001	0.100
	(0.0620)
Mar. 2001	0.144***
	(0.0555)
Apr. 2001	0.221***
	(0.0540)
May-01	0.191***
	(0.0523)

Jun. 2001	0.177***
	(0.0525)
Jul. 2001	0.166***
	(0.0528)
Aug. 2001	0.125**
	(0.0544)
Sep. 2001	0.168***
	(0.0590)
Oct. 2001	0.147***
	(0.0541)
Nov. 2001	0.117*
	(0.0669)
Dec. 2001	0.128**
	(0.0611)
Jan. 2002	0.116*
	(0.0596)
Feb. 2002	0.159***
	(0.0588)
Mar. 2002	0.215***
	(0.0542)
Apr. 2002	0.241***
	(0.0551)
May-02	0.240***
	(0.0557)
Jun. 2002	0.177***
	(0.0563)
Jul. 2002	0.211***
	(0.0500)
Aug. 2002	0.220***
	(0.0538)
Sep. 2002	0.263***
	(0.0525)
Oct. 2002	0.252***
	(0.0598)
Nov. 2002	0.255***
	(0.0533)
Dec. 2002	0.225***
	(0.0527)
Jan. 2003	0.281***
	(0.0528)

Feb. 2003	0.263***
	(0.0592)
Mar. 2003	0.265***
	(0.0524)
Apr. 2003	0.201***
	(0.0548)
May-03	0.246***
	(0.0535)
Jun. 2003	0.261***
	(0.0550)
Jul. 2003	0.274***
	(0.0539)
Aug. 2003	0.304***
	(0.0502)
Sep. 2003	0.263***
	(0.0518)
Oct. 2003	0.259***
	(0.0519)
Nov. 2003	0.249***
	(0.0586)
Dec. 2003	0.322***
	(0.0548)
Jan. 2004	0.210***
	(0.0610)
Feb. 2004	0.291***
	(0.0587)
Mar. 2004	0.262***
	(0.0583)
Apr. 2004	0.254***
	(0.0617)
May-04	0.338***
	(0.0513)
Jun. 2004	0.339***
	(0.0496)
Jul. 2004	0.302***
	(0.0501)
Aug. 2004	0.306***
	(0.0509)
Sep. 2004	0.349***
	(0.0532)

Oct. 2004	0.333***
	(0.0525)
Nov. 2004	0.345***
	(0.0544)
Dec. 2004	0.351***
	(0.0514)
Jan. 2005	0.271***
	(0.0724)
Feb. 2005	0.304***
	(0.0564)
Mar. 2005	0.350***
	(0.0495)
Apr. 2005	0.342***
	(0.0521)
May-05	0.310***
	(0.0508)
Jun. 2005	0.359***
	(0.0512)
Jul. 2005	0.362***
	(0.0499)
Aug. 2005	0.351***
	(0.0527)
Sep. 2005	0.362***
	(0.0527)
Oct. 2005	0.312***
	(0.0515)
Nov. 2005	0.351***
	(0.0549)
Dec. 2005	0.259***
	(0.0583)
Jan. 2006	0.289***
	(0.0553)
Feb. 2006	0.323***
	(0.0577)
Mar. 2006	0.360***
	(0.0503)
Apr. 2006	0.328***
	(0.0522)
May-06	0.342***
	(0.0505)

Jun. 2006	0.363***
	(0.0528)
Jul. 2006	0.323***
	(0.0550)
Aug. 2006	0.354***
	(0.0502)
Sep. 2006	0.316***
	(0.0525)
Oct. 2006	0.367***
	(0.0526)
Nov. 2006	0.323***
	(0.0540)
Dec. 2006	0.318***
	(0.0574)
Jan. 2007	0.323***
	(0.0544)
Feb. 2007	0.343***
	(0.0552)
Mar. 2007	0.317***
	(0.0557)
Apr. 2007	0.360***
	(0.0530)
May-07	0.310***
	(0.0507)
Jun. 2007	0.324***
	(0.0558)
Jul. 2007	0.295***
	(0.0521)
Aug. 2007	0.350***
	(0.0523)
Sep. 2007	0.335***
	(0.0584)
Oct. 2007	0.284***
	(0.0624)
Nov. 2007	0.305***
	(0.0574)
Dec. 2007	0.330***
	(0.0619)
Jan. 2008	0.254***
	(0.0575)

Feb. 2008	0.316***
	(0.0535)
Mar. 2008	0.339***
	(0.0534)
Apr. 2008	0.313***
	(0.0548)
May-08	0.303***
	(0.0557)
Jun. 2008	0.325***
	(0.0510)
Jul. 2008	0.243***
	(0.0554)
Aug. 2008	0.235***
	(0.0584)
Sep. 2008	0.264***
	(0.0590)
Oct. 2008	0.110
	(0.0878)
Nov. 2008	0.211***
	(0.0698)
Dec. 2008	0.249***
	(0.0744)
Jan. 2009	0.0989
	(0.113)
Feb. 2009	0.247***
	(0.0648)
Mar. 2009	0.213***
	(0.0630)
Apr. 2009	0.194***
	(0.0643)
May-09	0.266***
	(0.0582)
Jun. 2009	0.225***
	(0.0582)
Jul. 2009	0.262***
	(0.0537)
Aug. 2009	0.172***
	(0.0653)
Sep. 2009	0.125*
	(0.0688)

Oct. 2009	0.126**
	(0.0583)
Nov. 2009	0.195***
	(0.0623)
Dec. 2009	0.0412
	(0.0754)
Jan. 2010	0.105
	(0.0839)
Feb. 2010	0.0740
	(0.0819)
Mar. 2010	0.214***
	(0.0645)
Apr. 2010	0.189***
	(0.0562)
May-10	0.191***
	(0.0628)
Jun. 2010	0.193***
	(0.0582)
Jul. 2010	0.151**
	(0.0759)
Aug. 2010	0.207***
	(0.0722)
Sep. 2010	0.178***
	(0.0622)
Oct. 2010	0.149**
	(0.0658)
Nov. 2010	0.149**
	(0.0636)
Dec. 2010	0.0858
	(0.0711)
Jan. 2011	0.0866
	(0.0905)
Feb. 2011	0.0369
	(0.0801)
Mar. 2011	0.0746
	(0.0830)
Apr. 2011	0.213***
	(0.0705)
May-11	0.137*
	(0.0823)

Jun. 2011	0.233***
	(0.0629)
Jul. 2011	0.197***
	(0.0615)
Aug. 2011	0.137**
	(0.0657)
Sep. 2011	0.170**
	(0.0691)
Oct. 2011	0.249***
	(0.0679)
Nov. 2011	0.0670
	(0.0614)
Dec. 2011	0.199***
	(0.0709)
Jan. 2012	0.0893
	(0.0707)
Feb. 2012	0.166**
	(0.0819)
Mar. 2012	0.227***
	(0.0569)
Apr. 2012	0.169**
	(0.0685)
May-12	0.127**
	(0.0589)
Jun. 2012	0.199***
	(0.0641)
Jul. 2012	0.172***
	(0.0562)
Aug. 2012	0.160***
	(0.0553)
Sep. 2012	0.140**
	(0.0563)
Oct. 2012	0.158***
	(0.0564)
Nov. 2012	0.235***
	(0.0557)
Dec. 2012	0.0903
	(0.0636)
Constant	8.364***
	(0.154)

Observations	9,041
R-squared	0.749

Notes: robust standard errors in parentheses; the symbols *, **, and *** respectively denote statistical significance at the 10%, 5%, and 1% levels; the dependent variable is the log sale price of each home; census block group fixed effects are included in the model, but are not reported in this table.

Appendix 4. Modeling the effect of multiple discoveries on property values

This section presents results for the hedonic property models, where treatment is defined in terms of dosage. The hedonic price model is modified by taking the following form:

(A.4.1) $ln(P_{irt}) = \boldsymbol{X_{irt}}\beta_1 + \boldsymbol{Treated}_{irt}\beta_2 + \boldsymbol{TotalDisc}{*}After_{irt}\beta_3 + \boldsymbol{\tau}_t + \boldsymbol{\gamma}_r + \varepsilon_{irt.,}$

where the noteworthy term of distinction here is ***TotalDisc***. This term refers to the number of discoveries made within defined distance increments of property *i*, prior to its sale. The coefficient β_3 thus estimates the marginal effect of each discovery on a home's property value. For repeat sales, the model is similarly adjusted as follows:

(A.4.2) $\triangle ln(P_i) = \triangle \boldsymbol{TotalDisc}{*}After_i\beta_3 + \triangle\boldsymbol{\tau}_t + \triangle\varepsilon_i.$

When I consider the additional effect of public registers, I include an additional term which counts the number of such discoveries that, by my estimate, were disclosed by at least one public register at the time of the property's sale.

The results for the hedonic price model estimates are presented in Table A.4. As with the main results, in the interest of space only the results for the coefficient β_3 are reported. Generally speaking, these results suggest a negative treatment effect. In the case of columns (1) and (3), which do not differentiate between periods when registers are and are not accessible for reference, the effect declines with distance. The results of columns (3) and (4), which estimate the treatment effect of discoveries which are not classified as dump sites, are also generally similar to their counterparts in the first two columns. In the case of column (4), statistical significance is less directly apparent, suggesting a potentially nonlinear relationship between distance and the magnitude of the treatment effect. Note that, for sales which occurred when registers were available to document a given discovery, the marginal effect of such a discovery is calculated as the net value of the two coefficients, i.e. the “Total discoveries” and “Total discoveries disclosed via register”, for a defined distance increment. In the case of each of the two distance increments, the two variables are jointly significant.

Table A.5 presents results for the repeat sales model. These results are similar to those of the previous table, although here the relationship between the treatment effect and distance does not appear to be as linear, as suggested by overall results in columns (1) and (3). Columns (2) and (4) suggest that the treatment effect is most pronounced for properties affected by discoveries searchable by registers, at the time of sale. In the case of column (2), the “Total discoveries” and “Total discoveries disclosed via register” variables are jointly significant for each distance increment. In the case of column (4), neither distance increment produces jointly significance for these two variables. As noted in section 4.5, these results are likely to be sensitive to the small sample sizes employed for the repeat sales specifications.

Table A.4. Hedonic price model estimates, for homes within 400m of one or more discoveries

	(1)	(2)	(3)	(4)
	all meth discoveries		discoveries specifically denoted as "labs"	
Total discoveries, 200-400m	-0.0139***	-0.000178	-0.0155***	0.00241
	(0.00253)	(0.00299)	(0.00500)	(0.00645)
Total discoveries, 0-200m	-0.0216***	-0.00971*	-0.0218***	-0.00988
	(0.00460)	(0.00533)	(0.00774)	(0.0108)
Total discoveries, disclosed via register, 200-400m		-0.0344***		-0.0318***
		(0.00506)		(0.00896)
Total discoveries disclosed via register, 0-200m		-0.0278***		-0.0193
		(0.00915)		(0.0137)
Observations	26,231	26,231	20,001	20,001
R-squared	0.714	0.715	0.697	0.697

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A.5. Hedonic price model estimates, using repeat sales, for homes within 400m of one or more discoveries

	(1)	(2)	(3)	(4)
	all meth discoveries		discoveries specifically denoted as "labs"	
Total discoveries, 200-400m	-0.0101*	-6.66e-05	-0.00526	0.00650
	(0.00594)	(0.00688)	(0.00751)	(0.00867)
Total discoveries, 0-200m	-0.00357	0.00683	-0.00502	0.0108
	(0.00384)	(0.00428)	(0.0102)	(0.0142)
Total discoveries, disclosed via register, 200-400m		-0.0289**		-0.0232*
		(0.0125)		(0.0128)
Total discoveries disclosed via register, 0-200m		-0.0287***		-0.0281
		(0.00753)		(0.0188)
Observations	3,618	3,618	2,643	2,643
R-squared	0.416	0.421	0.430	0.432

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

References

Abbott, J.K., Klaiber, H.A., 2010. An embarrassment of riches: confronting omitted variable bias and multi-scale capitalization in hedonic price models. Review of Economics and Statistics, 93, 1331–1342.

Acharya, G., Bennett, L.L., 2001. Valuing open space and land-use patterns in urban watersheds. The Journal of Real Estate Finance and Economics, 22, 221–237.

Alberini, A., & Filippini, M., 2011. Response of residential electricity demand to price: The effect of measurement error. Energy Economics, 33(5), 889-895.

Alberini, A., and Towe, C., 2015. Information v. Energy Efficiency Incentives: Evidence from Residential Electricity Consumption in Maryland. CER-ETH–Center of Economic Research at ETH Zurich Working Paper, 15, 208.

Alberini, A., Gans, W., and Velez-Lopez, D., 2011. Residential consumption of gas and electricity in the US: The role of prices and income. Energy Economics, 33(5), 870-881.

Allcott, H. & Rogers, T., 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." American Economic Review, 104(10), 3003-3037.

Allcott, H., 2011. Social norms and energy conservation. Journal of Public Economics, 95(9), 1082-1095.

American Public Transportation Association, 2012. "Rural communities: expanding horizons." Available at: http://www.apta.com/resources/reportsandpublications/Documents/Rural-Communities-APTA-White-Paper.pdf. Accessed January 28, 2016.

Auffhammer, M., 2014. Quantifying Intensive and extensive margin adaptation responses to climate change: A study of California's Residential Electricity Consumption.

Baiocchi, G., Minx, J., & Hubacek, K., 2010. The impact of social factors and consumer behavior on carbon dioxide emissions in the United Kingdom. Journal of Industrial Ecology, 14(1), 50-72.

Bartos, L., 2005. "Meth production is toxic to communities." Available at: http://lancaster.unl.edu/family/methtoxic.shtml. Accessed August 27, 2016.

Benson, J., 2008. Questions and answers about methamphetamines. Available at: http://extension.missouri.edu/p/GH6201. Accessed August 27, 2016.

Bernard, J. T., Bolduc, D., & Yameogo, N. D., 2011. A pseudo-panel data model of household electricity demand. Resource and Energy Economics, 33(1), 315-325.

Bernstein, M. A., & Griffin, J. M., 2006. Regional differences in the price-elasticity of demand for energy. National Renewable Energy Laboratory.

Bin, O., Crawford, T. W., Kruse, J. B., & Landry, C. E., 2008. Viewscapes and flood hazard: Coastal housing market response to amenities and risk. Land Economics, 84(3), 434-448.

Black, S.E., 1999. Do better schools matter? Parental valuation of elementary education. The Quarterly Journal of Economics, 114, 577–599.

Blázquez, L., Boogen, N., and Filippini, M., 2013. Residential electricity demand in Spain: new empirical evidence using aggregate data. Energy economics, 36, 648-657.

Blundell, R., & Stoker, T. M. 2005. Heterogeneity and aggregation. Journal of Economic Literature, 43(2), 347-391.

Blundell, R., Pashardes, P., and Weber, G., 1993. What do we learn about consumer demand patterns from micro data? The American Economic Review, 83(3), 570-597.

Bockstael, N. E., & McConnell, K. E., 2007. Environmental and resource valuation with revealed preferences: A theoretical guide to empirical models (Vol. 7). Springer Science & Business Media.

Bohi, D. R., 1981. Analyzing demand behavior: a study of energy elasticities. Routledge.

Boogen, N., Datta, S., and Filippini, M., 2014. Going beyond tradition: Estimating residential electricity demand using an appliance index and energy services. CER-ETH–Center of Economic Research at ETH Zurich Working Paper, 14, 200.

Borenstein, S., 2009. To what electricity price do consumers respond? Residential demand elasticity under increasing-block pricing.

Bourguignon, F., Fournier, M., & Gurgand, M., 2007. Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. Journal of Economic Surveys, 21(1), 174-205.

Boxall, P.C., Chan, W.H., McMillan, M.L., 2005. The impact of oil and natural gas facilities on rural residential property values: a spatial hedonic analysis. Resource and Energy Economics, 27, 248–269.

Brasington, D.M., Hite, D., 2005. Demand for environmental quality: a spatial hedonic analysis. Regional Science and Urban Economics, 35, 57–82.

Büchs, M., & Schnepf, S. V., 2013. Who emits most? Associations between socio-economic factors and UK households' home energy, transport, indirect and total CO2 emissions. Ecological Economics, 90, 114-123.

Case, K.E. & Shiller, R.J., 1987. "Prices of single-family homes since 1970: new indexes for four cities," New England Economic Review, 45-56.

Chay, K.Y., Greenstone, M., 2005. Does air quality matter? Evidence from the housing market. Journal of Political Economy, 113, 376–424.

Commins, N., & Nolan, A., 2010. Car ownership and mode of transport to work in Ireland. The Economic and Social Review, 41(1), 43-75.

Congdon-Hohman, J.M., 2013. The lasting effects of crime: The relationship of discovered methamphetamine laboratories and home values. Regional Science and Urban Economics, 43, 31-41.

Cropper, M.L., Deck, L.B., Kishor, N.M., McConnell, K.E., 1993. Valuing product attributes using single market data: a comparison of hedonic and discrete choice approaches. The Review of Economics and Statistics, 75, 225–232.

Cropper, M.L., Deck, L.B., McConnell, K.E., 1988. On the choice of functional form for hedonic price functions. The Review of Economics and Statistics, 70, 668–675.

Dahl, G. B., 2002. Mobility and the returns to education: testing a Roy Model with multiple markets. Econometrica 70: 2367–2420.

Davis, L.W., 2004. The effect of health risk on housing values: evidence from a cancer cluster. American Economic Review, 94, 1693–1704.

Davis, L. W., 2008. Durable goods and residential demand for energy and water: evidence from a field trial. The RAND Journal of Economics, 39(2), 530-546.

Day, B., Bateman, I., Lake, I., 2007. Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic property price model. Environmental and Resource Economics, 37, pp.211–232.

Dergiades, T., and Tsoulfidis, L., 2008. Estimating residential demand for electricity in the United States, 1965–2006. Energy Economics, 30(5), 2722-2730.

Di Cosmo, V., and Hyland, M.,2013. Carbon tax scenarios and their effects on the Irish energy sector. Energy Policy, 59, 404-414.

Dewan, S., Brown, R., 2009, 'Illnesses afflict homes with a criminal past', The New York Times, 13 July, A1.

Dobkin, C., Nicosia, N., 2009. The war on drugs: methamphetamine, public health, and crime. American Economic Review, 99, 324–349.

Dresner, S., & Ekins, P., 2006. Economic instruments to improve UK home energy efficiency without negative social impacts*. Fiscal Studies, 27(1), 47-74.

Druckman, A., & Jackson, T., 2008. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. Energy policy, 36(8), 3177-3192.

Dubin, J. A., & McFadden, D. L., 1984. An econometric analysis of residential electric appliance holdings and consumption. Econometrica: Journal of the Econometric Society, 345-362.

Easter, M.G., 2010. "Are You Living in a Former Meth Lab?" Available at: http://www.scientificamerican.com/article/former-meth-lab/. Accessed March 6, 2016.

Emrath, P., & Liu, F., 2008. Vehicle carbon dioxide emissions and the compactness of residential development. Cityscape, 185-202.

Espey, J. A., & Espey, M., 2004. Turning on the lights: a meta-analysis of residential electricity demand elasticities. Journal of Agricultural and Applied Economics, 36(01), 65-81.

Fell, H., Li, S., & Paul, A., 2014. A new look at residential electricity demand using household expenditure data. International Journal of Industrial Organization, 33, 37-47.

Gamper-Rabindran, S., & Timmins, C., 2013. Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits. Journal of Environmental Economics and Management, 65(3), 345-360.

Gans, W., Alberini, A., and Longo, A., 2013. Smart meter devices and the effect of feedback on residential electricity consumption: Evidence from a natural experiment in Northern Ireland. Energy Economics, 36, 729-743.

Garcia-Cerrutti, L. M., 2000. Estimating elasticities of residential energy demand from panel county data using dynamic random variables models with heteroskedastic and correlated error terms. Resource and Energy economics, 22(4), 355-366.

Gately, D., and Huntington, H. G., 2002. The asymmetric effects of changes in price and income on energy and oil demand. The Energy Journal, 23(1), 19-55.

Gaur, V., and Gupta, E., 2016. The determinants of electricity theft: An empirical analysis of Indian states. Energy Policy, 93, 127-136.

Gawande, K., Jenkins-Smith, H., 2001. Nuclear waste transport and residential property values: estimating the effects of perceived risks. Journal of Environmental Economics and Management, 42, 207–233.

Geoghegan, J., Lynch, L., Bucholtz, S., 2003. Capitalization of open spaces into housing values and the residential property tax revenue impacts of agricultural easement programs. Agricultural and Resource Economics Review, 32, 33-45.

Gillingham, K., Jenn, A., & Azevedo, I. M. L., 2015. Heterogeneity in the response to gasoline prices: Evidence from Pennsylvania and implications for the rebound effect. Energy Economics, 52, Supplement 1, S41–S52.

Gillingham, K., Kotchen, M. J., Rapson, D. S., & Wagner, G., 2013. Energy policy: The rebound effect is overplayed. Nature, 493(7433), 475-476.

Glaeser, E. L., & Kahn, M. E., 2004. Sprawl and urban growth. Handbook of regional and urban economics, 4, 2481-2527.

Glaeser, E. L., & Kahn, M. E., 2010. The greenness of cities: carbon dioxide emissions and urban development. Journal of urban economics, 67(3), 404-418.

Goldberg, P. K., 1998. The effects of the corporate average fuel efficiency standards in the US. The Journal of Industrial Economics, 46(1), 1-33.

Gough, I., Abdallah, S., Johnson, V., Ryan-Collins, J., Smith, C., 2011. The distribution of total greenhouse gas emissions by households in the UK, and some implications for social policy CASE Paper 152, Centre for Analysis of Social Exclusion. London School of Economics, London.

Grainger, C. A., & Kolstad, C. D., 2010. Who pays a price on carbon? Environmental and Resource Economics, 46(3), 359-376.

Greenstone, M., Gallagher, J., 2008. Does hazardous waste matter? Evidence from the housing market and the Superfund program. The Quarterly Journal of Economics, 123, 951–1003.

Guignet, D., 2013. What do property values really tell us? a hedonic study of underground storage tanks. Land Economics, 89(2), 211-226.

Haas, R. & Schipper, L., 1998. Residential energy demand in OECD-countries and the role of irreversible efficiency improvements. Energy Economics, 20(4), 421–442.

Halvorsen, B., and Larsen, B. M., 2013. How serious is the aggregation problem? An empirical illustration. Applied Economics, 45(26), 3786-3794.

Hammar, Henrik, and Magnus Sjöström (2011), Accounting for Behavioral Effects of Increases in the Carbon Dioxide (CO_2) Tax in Revenue Estimation in Sweden, Energy Policy, 39, 6672-6676.

Hosoe, N., and Akiyama, S. I., 2009. Regional electric power demand elasticities of Japan's industrial and commercial sectors. Energy Policy, 37(11), 4313-4319.

Houthakker, H. S., 1980. Residential electricity revisited. The Energy Journal, 1(1), 29-41.

Hrenda, M., 2016. "DEU: Three Georgetown County residents charged with making, disposing of meth materials." Available at: http://www.southstrandnews.com/20160822/160829986/deu-three-georgetown-county-residents-charged-with-making-disposing-of-meth-materials. Accessed August 27, 2016.

Irwin, E.G., 2002. The Effects of Open Space on Residential Property Values. Land Economics, 78, 465–480.

Irwin, E.G., Bockstael, N.E., 2001. The problem of identifying land use spillovers: measuring the effects of open space on residential property values. American Journal of Agricultural Economics, 83, 698–704.

Ito, K., 2014. Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. American Economic Review, 104(2), 537-563.

Jamil, F., and Ahmad, E., 2014. An empirical study of electricity theft from electricity distribution companies in Pakistan. Pakistan Development Review, 53(3), 239-255.

Kamerschen, D. R., and Porter, D. V., 2004. The demand for residential, industrial and total electricity, 1973–1998. Energy Economics, 26(1), 87-100.

Katiyar, S. K., 2005. Political economy of electricity theft in rural areas: A case study from Rajasthan. Economic and Political Weekly, 644-648.

Kiel, K. A., & Williams, M., 2007. The impact of Superfund sites on local property values: Are all sites the same? Journal of Urban Economics, 61(1), 170-192.

Kiso, T., 2013. Essays on the Economics of Automobile Fuel Economy.

Krishnamurthy, C. K., and Kriström, B., 2013. Energy demand and income elasticity: a cross-country analysis. Available at SSRN 2310005.

Kuminoff, N.V., Parmeter, C.F., Pope, J.C., 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? Journal of Environmental Economics and Management, 60, 145–160.

Labandeira, X., Labeaga, J. M., and López-Otero, X., 2012. Estimation of elasticity price of electricity with incomplete information. Energy Economics, 34(3), 627-633.

Langer, A., and Miller, N. H., 2013. Automakers' Short-Run Responses to Changing Gasoline Prices. Review of Economics and Statistics, 95(4), 1198-1211.

Lee, L. F., 1983. Generalized econometric models with selectivity. Econometrica 51: 507–512.

Leggett, C.G., Bockstael, N.E., 2000. Evidence of the effects of water quality on residential land prices. Journal of Environmental Economics and Management, 39, 121–144

Lenzen, M., Wier, M., Cohen, C., Hayami, H., Pachauri, S., & Schaeffer, R., 2006. A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan. Energy, 31(2), 181-207.

Li, S., Linn, J., and Muehlegger, E.. 2014. Gasoline taxes and consumer behavior. American Economic Journal: Economic Policy, 6(4): 302-42.

Lim, K. M., Lim, S. Y., and Yoo, S. H., 2014. Short-and long-run elasticities of electricity demand in the Korean service sector. Energy Policy, 67, 517-521.

Linden, L., Rockoff, J., 2008. Estimates of the impact of crime risk on property values from Megan's laws. American Economic Review, 98, 1103–1127.

Maddala, G. S., Trost, R. P., Li, H., and Joutz, F., 1997. Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators. *Journal of Business & Economic Statistics*, *15*(1), 90-100.

Kahn, M. E., and Mansur, E. T., 2013. Do local energy prices and regulation affect the geographic concentration of employment?. *Journal of Public Economics*, *101*, 105-114.

Martin, D., 2014. The test homebuyers almost never do, but should. Available at: http://america.aljazeera.com/watch/shows/america-tonight/articles/2014/7/14/the-test-homebuyersalmostneverdobutshould.html. Accessed February 13, 2016.

McCluskey, J.J., Rausser, G.C., 2003. Stigmatized asset value: is it temporary or long-term? Review of Economics and Statistics, 85, 276–285.

Metcalf, G. E., and Hassett, K. A., 1999. Measuring the energy savings from home improvement investments: evidence from monthly billing data. *Review of economics and statistics*, 81(3), 516-528.

Mori, K., 2012. Modeling the impact of a carbon tax: A trial analysis for Washington State. Energy Policy, 48, 627-639.

Nakajima, T., and Hamori, S., 2010. Change in consumer sensitivity to electricity prices in response to retail deregulation: a panel empirical analysis of the residential demand for electricity in the United States. Energy policy, 38(5), 2470-2476.

Nakata, T., & Lamont, A., 2001. Analysis of the impacts of carbon taxes on energy systems in Japan. Energy Policy, 29(2), 159-166.

National Association of Counties, 2005. The meth epidemic in America: two surveys of US counties, 2005. Available at http://www.naco.org. Accessed June 23, 2013.

National Institute on Drug Abuse, 2013. “Research reports: Methamphetamine” Available at https://www.drugabuse.gov/publications/research-reports/methamphetamine. Accessed March 6, 2016.

Nelson, T., 2013. 7 Neighborhood Need-to-Knows for 21st Century Home Buyers. Available at: http://www.businessinsider.com/7-neighborhood-need-to-knows-for-21st-century-home-buyers-2013-12. Accessed Feburary 12, 2016.

Nechyba, T. J., & Walsh, R. P., 2004. Urban sprawl. Journal of economic perspectives, 177-200.

Nesbakken, R., 1999. Price sensitivity of residential energy consumption in Norway. Energy economics, 21(6), 493-515.

Nesbakken, R., 2001. Energy consumption for space heating: a discrete–continuous approach. The Scandinavian Journal of Economics, 103(1), 165-184.

Newcomer, A., Blumsack, S. A., Apt, J., Lave, L. B., & Morgan, M. G., 2008. Short run effects of a price on carbon dioxide emissions from US electric generators. Environmental Science & Technology, 42(9), 3139-3144.

Nutley, S., 2003. Indicators of transport and accessibility problems in rural Australia. Journal of transport geography, 11(1), 55-71.

Okajima, S., & Okajima, H., 2013. Estimation of Japanese price elasticities of residential electricity demand, 1990–2007. Energy Economics, 40, 433-440.

Palmquist, R.B., 2005. Chapter 16 Property Value Models. In Karl-Gran Mler and Jeffrey R. Vincent, ed. Handbook of Environmental Economics. Valuing Environmental Changes. Elsevier, pp. 763–819. Available at: http://www.sciencedirect.com/science/article/pii/S1574009905020164. Accessed September 9, 2013.

Parry, I. W., & Small, K. A., 2005. Does Britain or the United States have the right gasoline tax? The American Economic Review, 95(4), 1276-1289.

Paul, A. C., Myers, E. C., & Palmer, K. L., 2009. A partial adjustment model of US electricity demand by region, season, and sector.

Pope, D.G., Pope, J.C., 2012. Crime and property values: evidence from the 1990s crime drop. Regional Science and Urban Economics, 42, 177–188.

Pope, J.C., 2008a. Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise. Journal of Urban Economics, 63, 498–516.

Pope, J.C., 2008b. Fear of crime and housing prices: household reactions to sex offender registers. Journal of Urban Economics, 3, 601–614.

Quigley, J. M., and Rubinfeld, D. L., 1989. Unobservables in consumer choice: residential energy and the demand for comfort. The Review of Economics and Statistics, 71(3), 416-425.

Ready, R.C., Abdalla, C.W., 2005. The amenity and disamenity impacts of agriculture: estimates from a hedonic pricing model. American Journal of Agricultural Economics, 87, 314–326.

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

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy, 82, 34–55.

Ryan, D. L., Wang, Y., and Plourde, A., 1996. Asymmetric price responses of residential energy demand in Ontario. Canadian Journal of Economics, 29(1), S317-S323.

Ryan, L., Moarif, S., Levina, L., and Baron, R., 2011. Energy efficiency policy and carbon pricing: Information paper. Available at: https://www.iea.org/publications/freepublications/publication/EE_Carbon_Pricing.pdf. Accessed April 8, 2016.

Scott, M., 2002. Clandestine drug labs. Problem-oriented guides for police. U.S. Department of Justice. Available at: http://www.cops.usdoj.gov/pdf/e02021446.pdf. Accessed June 24, 2013.

Shin, J. S., 1985. Perception of price when price information is costly: evidence from residential electricity demand. The Review of Economics and Statistics, 591-598.

Small, K. A., & Van Dender, K., 2007. Fuel efficiency and motor vehicle travel: the declining rebound effect. The Energy Journal, 25-51.

Sorrell, S., & Dimitropoulos, J., 2008. The rebound effect: Microeconomic definitions, limitations and extensions. Ecological Economics, 65(3), 636-649.

Sorrell, S., 2007. "The rebound effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency". Available at: http://www.ukerc.ac.uk/programmes/technology-and-policy-assessment/the-rebound-effect-report.html. Accessed September 4, 2015.

Starkey, R., 2012. Personal carbon trading: A critical survey: Part 1: Equity. Ecological Economics, 73, 7-18.

Stoker, T. M., 1993. Empirical approaches to the problem of aggregation over individuals. Journal of Economic Literature, 31(4), 1827-1874.

Substance Abuse and Mental Health Services Administration, 2005. Results from the 2004 national survey on drug use and health: national findings. DHHS publication No. SMA 05-4062, NSDUH Series H-28. Rockville, MD.

The Countryside Agency; Department for Environment, Food and Rural Affairs; Office of the Deputy Prime Minister; Office for National Statistics; Welsh Assembly Government, 2004. Rural and urban classification 2004: An introductory guide. Available at: http://webarchive.nationalarchives.gov.uk/20130402151656/http://archive.defra.gov.uk/evidence/statistics/rural/documents/rural-defn/Rural_Urban_Introductory_Guide.pdf. Accessed March 1, 2016.

Thumim, J., & White, V., 2008. Distributional Impacts of Personal Carbon Trading: A report to the Department for Environment, Food and Rural Affairs. Defra, London.

U.S. Energy Information Administration (US EIA), 2014. "Price Elasticity for Energy Use in Buildings of the United States." Available at: http://www.eia.gov/analysis/studies/buildings/energyuse/pdf/price_elasticities.pdf. Accessed September 1, 2015.

US EIA, 2015. "Frequently Asked Questions: What is U.S. electricity generation by energy source?" Available at: https://www.eia.gov/tools/faqs/faq.cfm?id=427&t=3. Accessed February 11, 2016.

U.S. Environmental Protection Agency, 2011. "EPA Progress on the 2007 Methamphetamine Remediation Research Act." Available at: http://www.epa.gov/sites/production/files/2015-10/documents/20110927-11-p-0708.pdf. Accessed February 13, 2016.

U.S. Federal Highway Administration, 2012. "Planning for Transportation in Rural Areas: II. Our Rural Transportation System." Available at: http://www.fhwa.dot.gov/planning/publications/rural_areas_planning/page03.cfm. Accessed January 28, 2016.

Velaga, N. R., Beecroft, M., Nelson, J. D., Corsar, D., & Edwards, P., 2012. Transport poverty meets the digital divide: accessibility and connectivity in rural communities. Journal of Transport Geography, 21, 102-112.

Vinha, K.P., 2005. The impact of the Washington Metro on development patterns.

Weber, C., & Matthews, H., 2008. Quantifying the global and distributional aspects of American household carbon footprint. Ecological economics, 66(2), 379-391.

West, S. E., 2004. Distributional effects of alternative vehicle pollution control policies. Journal of Public Economics, 88(3), 735-757.

Zheng, S., Wang, R., Glaeser, E. L., & Kahn, M. E. (2010). The greenness of China: household carbon dioxide emissions and urban development. Journal of Economic Geography, 11(5), 761-792.