E2e Working Paper 024

Are Home Buyers Inattentive?
Evidence From Capitalization of Energy Costs

Erica Myers

Revised August 2017

This paper is part of the E2e Project Working Paper Series.

E2e is a joint initiative of the Energy Institute at Haas at the University of California, Berkeley, the Center for Energy and Environmental Policy Research (CEEPR) at the Massachusetts Institute of Technology, and the Energy Policy Institute at Chicago, University of Chicago. E2e is supported by a generous grant from The Alfred P. Sloan Foundation.

Are Home Buyers Inattentive? Evidence From Capitalization of Energy Costs

Erica Myers*

August 18, 2017

Abstract

This paper explores whether home buyers are inattentive to future energy costs. I exploit variation in energy costs in the form of fuel price changes in Massachusetts where there is significant overlap in the geographic and age distributions of oil-heated and gas-heated homes. I find that relative fuel price shifts cause relative changes in housing transaction prices consistent with full capitalization of the present value of future energy cost differences under relatively low discount rates. These findings are consistent with home buyers being attentive to energy costs and are not consistent with inattention.

Keywords: Inattention, Undervaluation, Myopia, Housing Prices, Energy Efficiency Gap

JEL Codes: D12, H25, R31

*Assistant Professor, Department of Agricultural and Consumer Economics, University of Illinois. Email: ecmyers@illinois.edu. I am grateful to Severin Borenstein, Maximilian Auffhammer, Meredith Fowlie, and Catherine Wolfram for their invaluable advice, and to Hunt Allcott, Judson Boomhower, Chris Bruegge, Benjamin Crost, Lucas Davis, Walter Graf, Koichiro Ito, Lala Ma and seminar participants at the EEE NBER Summer Institute, University of Chicago, and University of Illinois for helpful comments. Data support from the CoreLogic Academic Research Council (CLARC) Grant is gratefully acknowledged. I am also thankful for financial support from the Joseph L. Fisher Doctoral Dissertation Fellowship by Resources for the Future and from the Fisher Center for Real Estate & Urban Economics PhD Fellowship.

1 Introduction

Consumers are often more responsive to changes in purchase price than to less salient product costs such as shipping and handling expenses (Hossain and Morgan, 2006), sales tax (Chetty et al., 2009), and operating costs of appliances (e.g., Hausman, 1979). This type of consumer inattention has important implications for policy measures such as taxation, since in order to affect behavior, policies need to target costs to which people pay attention. In recent years, governments around the world have become interested in designing successful policy instruments for reducing greenhouse gas (GHG) emissions. Whether price based instruments such as taxes or cap-and-trade programs will be effective crucially depends on whether consumers are responsive to fuel prices in markets for energy-using durables. If people lack information about changes in energy prices or are inattentive to the resulting changes in the operating costs of energy using durables, they will under-invest in efficiency even under carbon pricing policies. In this case, other more traditional policy instruments, such as information campaigns, efficiency standards for appliances, or building codes, may be more effective at improving market efficiency.

This paper asks whether consumers are responsive to changes in energy prices in the housing market. The building sector is large contributor of U.S. GHGs. A growing proportion of annual emissions, around 40%, come from the residential and commercial buildings sector.[1] As end uses, space heating and cooling contribute almost as much to U.S. greenhouse gas emissions annually (13%) as personal vehicles (15%).[2] In recent years, consumers spent an average of $272 billion/year, on residential natural gas, electricity, and fuel oil, almost as much as they did on gasoline and motor oil.[3]

It can be challenging to estimate whether home buyers are attentive to energy costs.

[1]Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy, 2011 Buildings Energy Data Book (2012) pp. 1-1.

[2]EPA: Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2012 and EIA: Residential Energy Consumption Survey 2009, Commercial Buildings Energy Consumption Survey 2003.

[3]Bureau of Labor Statistics Consumer Expenditure Survey: Shares of annual aggregate expenditures and sources of income (2005-2014), inflated to 2014 USD.

Previous attempts have used hedonic approaches to see if utility bills (Johnson and Kaserman, 1983), measures of efficiency (Dinan and Miranowski, 1989), or efficiency letter grades (Brounen and Kok, 2011) are capitalized into housing sales. In general, these studies find that more efficient homes with lower fuel costs receive premiums in the housing market, which are consistent with relatively low implied discount rates. One limitation of this approach is that home efficiency is not randomly assigned, so that more efficient homes may be more likely to have renovations or other improvements that are unobservable to the researcher, but appreciated by home buyers. Therefore, the observed premium for efficient units may be due to unobserved differences in homes rather than the causal effect of energy cost savings.

This study is the first to estimate the effect of plausibly exogenous variation in energy costs on housing prices. I use changes in the relative fuel prices of heating oil and natural gas over time as a source of variation in energy costs. Natural gas is used to heat homes in most parts of the United States where substantial heating is required. However, in the northeastern United States 30-40% of households still heat with heating oil.[4] For this study, I focus on the state of Massachusetts, where there is significant overlap in the geographic and age distributions of oil-heated homes and gas-heated homes. I compare the transaction price of oil-heated versus gas-heated homes for the period 1990-2011, during which there is significant variation in the relative fuel prices. With two dominant heating fuels I am able to control for unobserved variation in the macroeconomic environment with year fixed effects. In addition, I observe multiple sales of homes, which allows me to control for time-invariant characteristics of a home with unit fixed effects. As the relative fuel price of a home increases, it should sell at a discount compared to homes with less expensive fuel. If home buyers are attentive to future fuel costs, this discount should reflect the net present value of the difference in fuel expenditure.

I find little evidence that home buyers are systematically "under-valuing" future fuel costs. When the relative cost of heating goes up by $1/MMBTU, it leads to a $3500 discount

[4]See American Housing Survey National Summary Table 2-5: Fuels-Occupied Units, years 2005, 2007, and 2009.

in relative housing transaction price. These results are consistent with full capitalization of the present value of fuel expenditure differences under a 3% discount rate .[5] It appears that home buyers are paying attention not only to whether a home heats with oil or gas, but the relative prices of those fuels and further, how those relative price differences translate into differences in the net present value of the future stream of payments.

These findings are relevant to the literature on consumer inattention in energy using durables, where evidence in other contexts has been mixed. Consistent with my findings, recent work in car markets suggests that consumers are relatively attentive to future fuel costs. Estimates of implied discount rates for automobile purchases range between 5% and 15% (Busse et al., 2013; Allcott and Wozny, 2014; Sallee et al., 2016; Grigolon et al., 2014). However, consumers might be less inattentive to gasoline purchases as opposed to residential fuels. Gasoline is one of the most salient products that consumers buy. Obtaining information about gasoline prices is relatively costless, since they are prominently posted at gas stations, and many people purchase gas one or more times in a week. In addition, people tend to know how much it would cost to drive from one place to another in car. Residential fuel costs may not be as well understood or salient for consumers. Households only receive bills on a monthly basis, making it harder to translate consumption of particular energy services into costs. In addition, consumers receive bills where natural gas is combined with electricity in many areas, potentially muddling individual effects.

Consistent with the possible role of salience, Myers (2017) finds that consumers are not attentive to fuel costs in rental housing markets. However, the decision to buy a home differs significantly from the decision to rent. Purchasing a home is one of the most significant investment decisions many people will make. Therefore, they may be more attentive to all aspects of a home purchase decision as opposed to a rental decision. In addition, home buyers differ from renters in important ways that may be correlated with attentiveness to energy costs, including age, family structure, financial stability, and employment status to

[5]As I describe in more detail in the empirical framework, I assume that consumers use a no-change forecast for future energy prices and I use an infinite time horizon for net present value estimates.

name a few.

Evidence on consumer attention has been varied in the context of appliances. Early work using a discrete choice framework found that consumers substantially discount future energy costs (e.g., Hausman, 1979; Dubin and McFadden, 1984). Rapson (2014), on the other hand, developed a structural model of air-conditioner demand and found that consumers value the stream of future savings from high efficiency units. Houde (2014) found that consumers are highly heterogeneous in how they value future energy prices in the context of refrigerators.

The finding that home buyers are paying attention to fuel prices and how those price movements translate into a stream of future cost differences suggests that fuel costs are well-understood and salient at the point of sale. This has important implications for carbon policy since an increasing proportion of U.S. carbon dioxide emissions come from the residential and commercial buildings sector. Because home buyers appear to be informed about and paying attention to fuel prices, pollution pricing will create incentives to reduce the amount of energy people choose to consume, to convert to cleaner heating fuels, and possibly to increase the efficiency of building shells and appliances.

This paper proceeds as follows. Section 2 describes the data, Section 3 details the empirical framework, Section 4 describes the results for the capitalization of energy costs into housing transaction prices, and Section 5 concludes.

2 Data

2.1 Housing Transaction and Characteristic Data

The real estate data firm, CoreLogic, provided the housing transaction and unit characteristic data with over 1,000,000 transactions in the state of Massachusetts between 1990 and 2011. The unit characteristic data contain information on the number of bedrooms, bathrooms, stories, square feet, year built, exterior wall type, heating fuel, and heating system type. The unit characteristic data and the transaction data were compiled by CoreLogic from different sources. As a result, the unit characteristic data provide one snapshot of a home and do not necessarily reflect the attributes at the point of sale. I carefully address this potential for measurement error in the empirical analysis.

Housing units were designated to be in one of 491 geographic units in order to protect the proprietary nature of the data. Each geographic unit is made up of 3-41 census block groups, with a mean size of 10 census block groups. The criteria used to group census block groups into geographic areas were (1) to allow no fewer than 10 sales within a geographic area in a year and (2) not to let the geographic areas cross natural gas utility or county boundaries. The larger geographic areas are less densely populated with fewer transactions.

I drop observations if a unit is sold more than once in a year, or more than 4 times over the 21 year sample period, indicating special circumstances such as foreclosure (about 13% of observations). In property records, the "effective age" of a building is adjusted for significant renovations or neglect. Over 99% of adjustments to property age in the sample were for improvements, so that the "effective year built" is later than the actual year built. I drop another 8% of the remaining sample for these types of large renovations or improvements. I use the middle 99% of the distribution of non-zero housing transactions, dropping the top and bottom 1/2% most extreme values. The remaining data used have 909,434 transactions with 604,807 housing units sold between 1 and 4 times. About 50% of the sample heats with oil and 50% heats with gas. Over half of the sample (60%) were sold only once during the sample period.

Massachusetts was chosen for this study because there is good geographic overlap between oil and gas houses. Figure 1 shows the proportion of oil homes by the geographic units described above. The white areas are Berkshire and Plymouth counties for which no transaction data were available. The darker areas represent geographies where a higher proportion of homes heat with gas. Very few of the geographies have less than 10% of homes heating with oil. This means that even where utility natural gas is available, there are still many houses that heat with oil. In western Massachusetts more homes are heated with oil because there is less population density, and in some areas, there is no utility gas available. Figure 2 displays which towns had utility natural gas service as of 2008.

Table 1 displays the results of t-tests comparing the means of the characteristics of oil and gas homes. Gas homes differ from oil homes in predictable ways. On average, gas homes are slightly younger, larger, and more expensive than oil homes. In addition, gas heating systems are most likely to use forced air, while oil heating systems are most likely to use forced hot water radiators. Figure 3 displays the distribution of the numbers of bedrooms, bathrooms, square feet, and year built for oil vs. gas units. Importantly, there is good overlap in the covariates between the two heating types, so there are good counterfactual comparisons in terms of characteristics as well as geographies.

2.2 Fuel Price Data

The natural gas price data are state-level average annual residential retail prices calculated as the consumption weighted average of state-level monthly prices reported by the EIA. The heating oil price data are the average annual New England (PADD 1A) number 2 heating oil residential retail prices calculated as the consumption weighted average of monthly prices reported by EIA. I inflated all prices to 2012 dollars using the consumer price index. Both natural gas and heating oil prices were converted into $/MMBTU in order to make them comparable. Figure 4 displays the price variation in residential natural gas and heating oil prices from 1990 to 2012. In the mid-1990s, heating oil was less expensive than natural gas. But, starting in the mid-2000s, the price of heating oil began to rise, driven by world oil demand. The price of natural gas was rising in the early 2000s, until the use of hydraulic fracturing techniques began to drive prices down after 2006. Figure 5 shows the price difference (price of oil-price of gas) between the two fuels over the time period. Importantly, the price difference follows a "U" shape rather than a simple linear trend allowing me to identify the effects of fuel price variation rather than other trending variables on housing prices.

3 Empirical Framework

A general test researchers have used for systematic misperception of potentially less salient costs such as shipping and handling, sales tax, or automatic electronic payments is to compare the demand response of those costs versus salient, correctly perceived costs (Chetty et al., 2009; Hossain and Morgan, 2006; Finkelstein, 2009). Researchers have applied this test to energy using durables, comparing demand response to potentially misperceived future energy costs versus upfront purchase costs (Allcott and Wozny, 2014; Dubin and McFadden, 1984; Goldberg, 1998; Grigolon et al., 2014; Houde and Spurlock, 2015; Hausman, 1979). The intuition is that consumers should be indifferent between an additional dollar of purchase price and an additional present discounted dollar of energy expenditure, since total lifetime cost should be the relevant metric.

3.1 Theoretical Framework

In this paper I take a similar approach, estimating how home buyers' tradeoff purchase price and energy costs. In what follows, I develop a discrete choice framework where home buyer (i) chooses house (j) in geographic area (a) in year (t) from a choice set with budget constraint w_i. The consumer has an outside option of not buying a house with a utility level normalized to zero. Consumer i's indirect utility from the purchase of a home is a function of the cost, which has two components: 1) the transaction price, H_{jat}, and 2) the net present value (NPV) of the expected stream of future fuel payments, F_{jat}. Utility is also a function of observable home attributes, $\mathbf{X}_{jat}$, unobservable home attributes, $\tilde{\xi}_{jat}$, neighborhood-year specific amenities, $\tilde{\lambda}_{at}$, and individual taste ε_{ijat} as follows.

$$U_{ijat} = \eta(w_i - H_{jat} - \gamma F_{jat}) + \mathbf{X}'_{ja}\tilde{\beta} + \tilde{\xi}_{ja} + \tilde{\lambda}_{at} + \varepsilon_{ijat} \tag{1}$$

The marginal utility of money is represented by η. The implied discount rate is the discount rate that consumers would have to be using for $\gamma = 1$. If the implied discount rate is higher than the borrowing rate for the marginal dollar, then consumers are inattentive energy costs.[6] In other words, demand for homes with high fuel costs is too high relative to what would be optimal.

The choice to sell or buy a home in any given year is driven by exogenous events such as changes in employment or changes in family composition. All potential home buyers in geographic area a in year t have the same income and face the same choice set of homes, where they are trading off the price of a home versus attributes such as square footage or number of bedrooms and bathrooms. Assume a traditional representative consumer logit model where ε_{ijat} is distributed i.i.d. extreme value. Integrating over ε_{ijat} and taking the natural log of both sides gives the following relative choice probability as a function of prices and characteristics.

$$\frac{1}{\eta}(ln\phi_{jat} - ln\phi_{0at}) = H_{jat} - \gamma F_{jat} + \mathbf{X}'_{ja}\beta + \xi_{ja} + \lambda_{at} \tag{2}$$

On the left hand side is the choice probability for a house, ϕ_{jat}, relative to the choice probability of the outside option, ϕ_{0at}. Dividing by η gives the new variables $\mathbf{X}'_{ja}\beta$, ξ_{ja}, and δ_{at}, which can be interpreted as dollar value of the utility represented by $\mathbf{X}'_{ja}\tilde{\beta}$, $\tilde{\xi}_{jt}$, and $\tilde{\delta}_{at}$.

This can be rearranged into an econometric estimating equation of transaction price as a function of fuel costs and a set of fixed effects as follows:

$$H_{jat} = \gamma F_{jat} + \lambda_{at} + \theta_{ja} + \epsilon_{jat} \tag{3}$$

Variation in the probability of choosing the outside option over time and across space is

[6] The word "inattentive" means that consumers are undervaluing a dollar spent on future energy costs relative to a dollar spent in upfront price. This "mistake" might arise through several potential mechanisms such as imperfect information, biased beliefs, present bias, or bias toward concentration.

absorbed by geographic area by year fixed effects λ_{at}, which also control for shocks common to all houses in a given geographic area in a given year. House specific fixed effects (θ_{ja}) control for time invariant observable (X_{ja}) and unobservable characteristics (ξ_{ja}). Since the same house is sold more than once, it is being perceived by different sets of buyers across time periods. The new error term $\epsilon_{jat} = ln\phi_{jat}$ represents the idiosyncratic changes in the preferences for particular house due to the buyers in a particular period, and is uncorrelated with fuel price. This is a similar theoretical approach to that taken in Allcott and Wozny (2014) in the context of car markets, which uses cross-sectional variation in fuel economy interacted with variation over time in gasoline prices to get plausibly exogenous variation in lifetime fuel costs of cars. In this analysis, I use relative fuel price movements of oil and natural gas as a plausibly exogenous instrument for F_{jat}.

3.2 Estimation

One of the primary challenges of this type of exercise in the context of energy using durables is that we do not observe F_{jat} nor its underlying parameters directly. The NPV of the stream of expected future fuel payments, F_{jat}, is a function of the relevant time horizon (T), the discount factor (δ^i), the expected future fuel prices (p_{jai}), and expected future energy consumption (e_{jai}), where i indexes future years.

$$F_{jat} = \sum_{i=t}^{T} \delta^i p_{jai} e_{jai} \tag{4}$$

One approach researchers have taken is to estimate Equation 3 in a two stage process (e.g., Allcott and Wozny, 2014; Sallee et al., 2016; Grigolon et al., 2014). In the first stage, they calculate F_{jat} directly, making assumptions about T, e_{jai}, the correct discount factor, δ^i, and consumers beliefs about p_{jai}. Then, Equation 3 is estimated as a second stage. If a researcher can credibly calculate NPV of the stream of expected future fuel payments F_{jat}, this approach has the advantage of ease of interpretation. The coefficient estimate γ is directly interpretable as a measure of inattention: the demand responsiveness to the NPV of a dollar spent on energy in future years relative to a dollar spent on upfront price. Another advantage is that researchers can cleanly incorporate heterogeneity in F_{jat}, by calculating house-specific fuel costs that may vary by square-footage or other characteristics.

However there are disadvantages to this approach in my empirical setting, affecting the practical ability to estimate Equation 3 using the two stage approach. Since I do not directly observe billing data for each house, estimating F_{jat} as a function of house characteristics as a first stage has the potential to introduce bias in the second stage. If the second stage included any variables that affect energy expenditure, and these variables are not included in the first stage, it would introduce mechanical correlation between the first stage residuals and those variables only included in the second stage.[7] This is the same argument as to why the same exogenous covariates need to be included in both the first and second stage in any two stage least squares (2SLS) estimation (Wooldridge, 2010).

For example, residential energy consumption surveys (RECS) provide a repeated cross-section of energy expenditure, energy consumption, heating fuel type, and housing characteristics. However locational information is limited to large areas such as census regions, which are aggregations of several states. Therefore, it is not possible to estimate the first stage with either house fixed effects or geographic area-by-time fixed effects. If the fixed effects are included in the second stage, but not the first stage, they will likely be correlated with the first stage residuals since unobservable factors that affect housing price likely also affect energy consumption. The inconsistency from the correlation can spillover to all coefficient estimates in the second stage .[8]

Therefore, I begin by estimating the reduced form for my fuel price instrument and

[7] A simulation exercise demonstrating the empirical importance of the bias in this setting can be provided upon request.

[8] For completeness, I provide two-stage estimates in Appendix A3. However, it is not possible to control for location-specific trends or unit fixed effects, which affects the interpretation of the estimates. I discuss the estimation and results in the appendix of the paper.

interpret the implied discount rate of these coefficients *ex post*, similar to Busse et al. (2013). The advantage of starting with the reduced form is that the coefficient $\tilde{\gamma}$ is readily interpretable as the effect of contemporaneous fuel price movements on housing prices. It is also straightforward to use this estimate to calculate implied discount rates in Equation 3 given assumptions about F_{jat}, and there is a direct mapping to γ. In addition, without the limitations of the first stage estimation, I can incorporate the full range of fixed effects and perform an instrumental variables strategy for measurement error in heating fuel type in the CoreLogic data, which I describe in detail below. The estimation equation for the reduced form is as follows.

$$H_{jat} = \tilde{\gamma} p_{jt} + \lambda_{at} + \theta_{ja} + \epsilon_{jat} \tag{5}$$

The fuel price, p_{jt}, is the annual residential retail fuel price for Massachusetts and varies by whether house j is oil or gas heated. I use one statewide average price for each of these fuels, since more localized price variation may introduce endogeneity if, for instance, utility rates change coincident with some other local market factor affecting housing price.

The coefficient $\tilde{\gamma}$, is the effect of a \$1/MMBTU heating fuel price increase on the housing transaction price. As the relative fuel price of a home increases, it should sell at a discount compared to homes with less expensive fuel. If home buyers correctly perceive future fuel costs, this discount should reflect the change in F_{jat} caused by a \$1/MMBTU heating fuel price increase.

The advantage of having two different primary heating fuels is that I can separate the effect of fuel price movements on housing prices from the effect of trends in other macroeconomic variables. Since there is no cross-sectional variation in fuel price, one fuel price is collinear with year fixed effects, so that the identifying variation of the instrument is the difference between the price of oil and the price of natural gas from year to year. The identifying assumption for this approach is that oil and gas houses do not systematically differ in an unobservable or inadequately controlled for way that is correlated with the difference in fuel price.[9]

In what follows, I describe my approach for estimating the change in F_{jat} caused by a \$1/MMBTU heating fuel price increase.

3.2.1 Time Horizon

Houses are long-lived assets with some houses in the sample being over 300 years old. Because the assets are so long-lived, the correct time horizon to consider for the flow of future energy costs, F_{jat}, could potentially be infinite. For this analysis, I provide estimates based on an infinite time horizon, which is a conservative benchmark. If consumers were truly considering a shorter time horizon, assuming an infinite time horizon would lead to higher estimates of implied discount rates, and bias my analysis toward finding consumer inattention.

3.2.2 Beliefs about Future Fuel Prices

For my main specifications, I assume that consumers believe that annual fuel prices follow a no-change forecast, so that contemporaneous annual fuel prices are the best predictor of future annual fuel prices. A recent study by Anderson et al. (2013) finds that consumers believe that gasoline prices follow this type of pattern. In the case of heating oil, a no-change forecast predicts future crude oil prices as well as or better than forecasts derived from futures markets or surveyed experts (Alquist and Kilian, 2010; Alquist et al., 2012).

Another possibility is that consumers use information from crude oil and natural gas futures markets to make projections about fuel prices going forward. Figure 6 shows the

[9] Another, equivalent way to set up the reduced form would be to regress sale price on the price difference between oil and natural gas interacted with indicators for the type of fuel used to heat the house:

$$H_{jat} = \beta_0 + \beta_1 \mathrm{I}^{\text{oil}}_{jat} \times (p^{\text{oil}}_t - p^{\text{gas}}_t) + \beta_2 \mathrm{I}^{\text{gas}}_{jat} \times (p^{\text{oil}}_t - p^{\text{gas}}_t) + \lambda_{at} + \theta_{ja} + \epsilon_{jat}$$

β_2 drops out of the estimation, since it is collinear with year fixed effects. The estimate of β_1 is equivalent to that in Equation 5. I derive the equivalence of these two approaches in Appendix A1 spot and forward curves for crude oil (panel A) and natural gas (panel B). The natural gas forward curves reflect seasonality in prices, whereas the crude oil forward curves are much smoother. Panel C of Figure 6 shows the difference in the spot and forward prices between the two fuels (price of oil - price of gas).

One thing to note about the relationship between the spot and future curves of these two fuels is that the forward curves do not deviate substantially from spot prices. Therefore, even if home buyers were and paying attention to trends in futures prices, their beliefs about fuel prices going forward would not differ significantly from no-change beliefs.[10]

Given the long-run time horizon, consumers could expect reversion to some "long run" fuel price. A forecast based on this type of mean reverting process would mean that the contemporaneous price would overestimate the true perceived price difference in years with high differentials and would underestimate the true perceived price difference in years with low differentials. Therefore, if I assume consumers are using contemporaneous prices when they are actually expecting a reversion to the long run mean, it will bias estimates toward finding undervaluation of energy costs. It is also not clear how to implement an estimation assuming long run mean reversion, since consumers expectations long future prices would not vary on short-run basis.

3.2.3 Future Energy Consumption

If consumers believe that 1) future consumption will be a function of future fuel prices and 2) the best predictor of future fuel prices are today's fuel prices, then it is reasonable for them to believe that future consumption will be similar to today's consumption. If so, a good approximation of the change in the NPV fuel expenditure (F_{jat}) caused by a small changes in today's price would be the sum of the present value of the change in today's price times the discounted present value of today's quantity summed over each year in the relevant time

[10] I test this assertion more rigorously by using the discount-factor weighted average futures price rather than contemporaneous price for the analysis. I discuss the estimation procedure and provide resultsm in Appendix A4

horizon.

$$\frac{\partial F_{jat}}{\partial P_{jt}} = \left(\sum_{i=t}^{T} \delta^i e_{jat} \right) \cdot \Delta p_{jt} \tag{6}$$

For larger price differences, home owners may consume less energy and have a lower NPV of future fuel expenditure as a result of this elasticity. However, as I demonstrate below, demand for residential energy tends to be relatively inelastic, so incorporating reasonable estimates for demand elasticity changes expected future expenditure little.

I approximate consumption for the mean household using the value reported by the Residential Energy Consumption Survey (RECS) for single family homes in the northeast census region that heat with oil or gas, 94 MMBTU per year.[11]

If homes were perfectly inelastic, a \$1/MMBTU price difference would result in an annual reduction in energy expenditure of \$94 per year. Changing the assumed elasticity from zero to the higher end of residential fuel elasticity estimates, -.6, lowers the estimated annual change in expenditure from a \$1/MMBTU fuel price increase a small amount from \$94/year to \$90/year.[12]. I assume a locally linear demand curve and the mean fuel price in the sample, \$14.67/MMBTU, to calculate the change in consumption due to change in price. In the analysis, I will take the middle of this range, \$92, to calculate implied discount rates.

3.2.4 Estimating Implied Discount Rates

Given assumptions about: 1) beliefs about future fuel prices, 2) the relevant time horizon, and 3) future energy consumption, it is straightforward to estimate an implied discount rate given the reduced form estimation in Equation 5.

The coefficient $\tilde{\gamma}$ is the effect of a change in fuel price on the housing transaction price.

[11]Residential Energy Consumption Survey: Table CE2.2 "Household Site Fuel Consumption in the Northeast Region, Totals and Averages, 2009."

[12]This range of elasticity estimates is consistent with demand for residential electricity (e.g., Reiss and White, 2008; Ito, 2014; Jessoe and Rapson, 2014)

If consumers correctly perceive future fuel costs, $\tilde{\gamma}$ should be equal to the change in the NPV of future energy cost caused by the change in price ($\frac{\partial F_{jat}}{\partial P_{jt}}$). The implied discount rate is the discount rate that consumers would have to use for this to be true and the following to hold.

$$1 = \frac{\tilde{\gamma}}{\frac{\partial F_{jt}}{\partial p_{jt}}} \tag{7}$$

3.3 Cost of Conversion from Oil to Natural Gas

If the price of oil gets high enough compared to natural gas, it could be the case that the net present value of fuel expenditure difference between heating with oil and heating with natural gas exceeds the typical cost of conversion. In that case, economic theory would predict that the housing transaction price differential would not exceed the cost of conversion.

The cost of converting from oil to gas can vary widely from a few thousand dollars to over $10,000 (Notte, 2012). The cost of conversion depends on several factors including the system you choose to install, whether or not you have an underground oil tank that needs to be removed, and the cost of connecting to the main supply line. Conversion can be much more costly in areas that do not have access to the main supply line for natural gas. In many cases, utilities will extend the supply line only if residents are willing to pay for it.

If the conversion cost ceiling were a large biasing factor in this analysis, the cost of conversion would act as a limit on the level of pass-through of the expenditure differential, particularly in later years when the fuel price difference is large. As I show in the results section, this does not appear to a major concern, since later years have similar implied discount rates to earlier years.

4 Results

4.1 Basic Specification

In this section, I estimate the reduced form effect of relative fuel price shifts on relative transaction price and calculate the implied discount rates from the estimates. My preferred specification includes house fixed effects and geographic area by time fixed effects as in Equation 5. As the relative fuel price of a home increases, it should sell at a discount compared to homes with less expensive fuel, reflecting the net present value of the difference in fuel expenditure.

Table 2 displays the results from the estimation of the preferred specification (column 5) as well as several models with less flexible controls (columns 1-4). The first two columns show estimates for a model that includes year fixed effects with and without housing attribute controls. Housing attribute controls include flexible controls for decade built, number of stories, number of bedrooms, number of bathrooms, exterior wall type, heating system type, and square footage binned for every 500 square feet for unit i. I_i^{oil} indicates whether unit i heats with oil. The estimates in columns 3-4 come from models with geographic area by year fixed effects and housing unit fixed effects respectively. The estimates in column 5 for the preferred specification include both geographic area by year and housing unit fixed effects. Robust standard errors are two-way clustered at the house and geographic area by year levels to account for both autocorrelation between sales and correlation due to geographic area-specific shocks.

The results indicate that when the relative cost of heating goes up by \$1/MMBTU, it leads to a \$1000-\$1200 discount in relative housing transaction price. The last row of the table shows the implied discount rate for the coefficient estimate, assuming a increase in annual energy expenditure of \$92 per \$1/MMBTU increase in fuel price over an infinite time horizon. The results imply that home buyers use a 8-10% discount rate, which suggests that they do not strongly under-value future heating fuel costs when purchasing houses. For instance, these results are consistent with recent work on automobile purchases that also find no evidence of strong undervaluation. Recent estimates of implied discount rates for automobile purchases range between 5% and 15% (Busse et al., 2013; Allcott and Wozny, 2014; Sallee et al., 2016; Grigolon et al., 2014)

In addition, the estimation procedure does not appear to be sensitive to the particular energy usage assumptions nor the use of contemporaneous rather than futures prices. The results are quite close to those from a two-stage procedure with a limited set of housing characteristics and fixed effects, with an implied discount rate of 10%. This suggests that the results using a single estimated annual usage are a good approximation of the implied discount rate for the average home in the sample.[13] The futures prices are close enough to spot prices that estimates using a discount-factor weighted average of futures prices rather than contemporaneous prices yield the same implied discount rate as the preferred specification (9.5%).[14]

Figure 7 displays the relationship between the housing transaction price difference for oil versus gas homes and the the net present value of the difference in annual expenditure from heating with gas as opposed to oil. The left side of Figure 7 plots this relationship over the sample period. I estimate the NPV of the difference in fuel expenditure between heating with oil and gas over an infinite horizon using the estimate of a change in annual expenditure of $92 per $1 difference in relative fuel price and a 9.5% discount rate from the preferred estimation in Table 2. In addition, I depict the housing transaction price difference between gas and oil homes from the preferred specification with geographic area by year fixed effects and unit fixed effects by plotting coefficients on the year-specific gas intercepts ($\beta_1 - \beta_{22}$) from the following regression.

$$\mathrm{H}_{jat} = \sum_{t=1}^{22} \beta_t \mathrm{I}_{ja}^{\mathrm{gas}} + \lambda_{at} + \theta_{ja} + \epsilon_{jat} \tag{8}$$

[13]The data used and a full description of the two-stage estimation can be found in Appendix A2 and Appendix A3 respectively.

[14]Appendix A4 describes this estimation in detail.

The housing transaction price H_{jat} for house j in geographic area a in year t is regressed on house fixed effects, δ_{ja}, geographic are by time fixed effects, γ_{at}, and year-specific gas intercept terms where $\mathrm{I}_{ja}^{\mathrm{gas}}$ indicates the home heats with gas and ϵ_{jat} is the idiosyncratic error term. In the left side of Figure 7, the variation in housing price difference tracks the NPV of the difference in expenditure closely over the sample period.

The right side of Figure 7 plots the fuel price difference against the corresponding NPV of the difference in fuel expenditure for each year in the sample. If the housing transaction price difference was precisely the estimated NPV of the difference in fuel expenditure, each dot would fall on the 45 degree line. The fitted line through the scatter plot shows that the NPV estimate of the fuel expenditure difference using a 9.5% discount rate is a close fit for the housing transaction price difference.

4.2 Addressing Differential Trends and Measurement Error

One potential worry with this approach is that the pattern in relative housing transaction prices is caused by a differential trend in homes with a particular heating fuel rather than by the relative fuel price variation. For example, since oil homes are older on average, the results might be explained by the declining value of a vintage over time. In other words, when oil is getting most expensive relative to natural gas in later years, oil homes are also getting older on average compared to natural gas homes. This trend in age difference might partially explain some of the observed discount for oil homes compared to natural gas homes.

In order to address the issues of differential trends, columns 1 and 2 of Table 3 display results for two additional controls. First, for the estimates in column 1, I included an oil-heat linear trend. If my results were the result of a differential trend in homes that heat with oil rather than fuel price variation, the inclusion of the trend would substantially change the estimates. While the estimates do not change substantially, they are somewhat attenuated.

Second, for the estimates in column 2, I flexibly control for the age of the home with age fixed effects where age is defined by the sales year minus year built. Age fixed effects allow me to control flexibly for trends in value of houses as they age. Homes are grouped in 5 year bins for homes that are 20 years or older, because there are relatively few observations for each vintage in early years. For homes younger than 20 years, I use the actual year built, because there are more observations for each vintage year and the value of newer homes is likely much more sensitive to smaller age differences. Here too, the coefficients of interest are somewhat attenuated, suggesting these trends may drive some of the observed variation in housing price.

Another potential concern with my approach is the measurement error introduced by the housing unit characteristic data. As is the case with most real estate transaction data, the unit characteristic data provide only a snap shot of a house's attributes even though the transaction data span over 20 years. Therefore, there is a potential for measurement error in the characteristics at the point of sale. Measurement error, particularly in the heating fuel, could potentially bias the estimates.

If the measurement error were classical, it would attenuate the estimates toward zero and make it more likely to find evidence consistent with undervaluation of energy costs. However, in this context it is likely that the measurement error is non-classical. The more recent housing transactions are more likely than earlier transactions to have the correct housing characteristics. In later years as the price of oil increases compared to natural gas, people maybe converting from oil to gas. This has the potential to bias the estimates toward finding high levels of capitalization and away from finding myopia. The intuition is that in early years, when there is more likely to be measurement error, the estimate of the mean difference in housing transaction prices is more likely to be attenuated, while in later years, the difference in housing transaction price is likely to be more precise. Since the biggest change in fuel prices is in later years, some of the difference in housing transaction price attributed to change in fuel price may be driven instead by the increasing precision of the estimates.

Another source of potential bias stems from the fact that homeowners may be improving other aspects of the home that are unobserved in the data when they are changing heating fuel. For example they may choose to put in new flooring or new kitchen appliances such as a gas stove. Then houses may have an unobservably higher quality after they convert than before. If conversions are correlated with the price difference and are accompanied by other major renovations, it will exacerbate the non-classical measurement error problem, biasing the estimates away from zero, making it more likely to find evidence of capitalization.[15]

In order to address this issue and the issue of non-classical measurement error while controlling for trends in oil and vintage, I consider an instrumental variables approach, creating an instrument for heating fuel. I exploit temporal variation in the fuel type of new construction in order to isolate variation in fuel choice that is exogenous to the fuel price difference. Figure 8 displays the proportion of homes in the sample built with oil for each vintage year from 1900 to 2011. It is clear that there is variation in fuel choice that is separable from a linear trend in vintage. Figure 8 depicts several clear breaks in the data, which are associated with policy changes that are exogenous to the local housing market.

In late 1953, piped natural gas began to be delivered to New England. Prior to 1953, the region almost exclusively used manufactured gas (Castaneda, 1993). There is a sharp kink in the proportion of homes built with oil starting in 1953. After 1953, more and more homes are built with gas until about 1974. The price control policy lead to shortages in supply in the mid-1970's. The way that many utilities dealt with these shortages was to restrict access to new customers rather than by rationing existing consumers (Davis and Kilian, 2011). Between 1974 and 1978, there was an increase in homes built with oil due to supply shortages of natural gas. In 1978, wellhead price controls were lifted, which increased natural gas supply, and the proportion of homes built with oil dropped. Since then, natural gas has been getting more common with the exception of a brief increase in homes built with oil in the mid-1980's following the crude oil price collapse of 1986.

[15] As stated in the data section, in the initial data construction, I did remove any houses that appear to have had major upgrades, possibly reducing the prevalence of homes with major endogenous upgrades. However the instrumental variables approach addresses potential bias arising from their presence in the sample.

Using this variation, the instrument for price is the proportion of homes built with oil in the year a particular house was built times the oil price in the year it was sold plus the proportion of homes built with gas in the year a particular house was built times the gas price in the year it was sold. Using this instrument, the local average treatment effect will come from a comparison of vintages when gas was more or less readily available. I include an oil specific trend and flexibly control for the age of the house.[16]

The results from this estimation are displayed in Table 4. The first column shows the results of the first stage estimation. The coefficient on the instrument is close to one since on average, the instrument closely predicts fuel price. Column 2 shows the results of the two-staged least squares estimation. The point estimate of the price coefficient using two-staged least squares is much larger than that using OLS. This suggests that the measurement error, even though it is non-classical, may have served to attenuate rather than bias the estimates upward. The implied discount rate for the two-staged least squared estimate is around 3%, close to recent mortgage interest rates and the 95% confidence interval (1.7 - 6.3%) rules out values as high as those suggested by the basic OLS estimation.

4.3 Placebo Test

The reduced form IV regression compares the relative housing prices for vintages with a high proportion of oil-heated homes to vintages with a low proportion of oil-heated homes as oil prices move relative to gas prices, controlling for age, house fixed effects and geographic area by time fixed effects. The exclusion restriction therefore requires that, conditional on the controls, the only way the proportion of oil built in a particular year affects housing price is through fuel price. In what follows, I perform a placebo test to probe the validity of this assumption.

[16] As with the basic specification, an equivalent approach would be to estimate

$$H_{jat} = \beta_0 + \beta_1 \mathrm{I}^{\mathrm{oil}}_{jat} \times (p^{\mathrm{oil}}_t - p^{\mathrm{gas}}_t) + \lambda_{at} + \theta_{ja} + \epsilon_{jat}$$

where the proportion of homes built with oil in the year a particular house was built times the price difference between oil and gas is the instrument for oil times the price difference between oil and natural gas.

I remove gas-heated homes from the sample and perform the IV estimate on just oil heated homes. I define oil homes built in years where more than half of homes built were oil-heated as "placebo oil" and oil homes built in years where fewer than half of homes built were oil-heated as "placebo gas" The reduced form is exactly the same as before, the only difference being gas homes have been removed. If the exclusion restriction did not hold and other unobserved or inadequately controlled for characteristics about vintages with a high/low proportion oil heating were driving the results, then the placebo test should yield negative and significant results as in Table 4.

Table 5 displays the results from this placebo test. The point estimate is more than an order of magnitude lower and is statistically indistinguishable from zero. This suggests that the exclusion restriction holds, and the capitalization rates in Table 4 are driven by fuel price variation and not by unobservable or inadequately controlled for trends in vintage.

5 Conclusion

This paper explores how shifts in energy costs affect housing transaction prices to see if home buyers are inattentive to energy costs. I use shifts in natural gas and heating oil prices over time to isolate exogenous variation in home energy costs. I use housing transaction data from Massachusetts, where roughly an equal number of homes heat with oil as heat with natural gas. This allows me to estimate the effect of a change in relative energy costs on a change in relative housing prices, while controlling for changes in the macroeconomic environment and in the value of different housing characteristics over time.

I find that home buyers are relatively attentive to future fuel costs. They are paying attention to shifts in relative fuel prices and are aware of how changes in fuel price translate into changes in the net present value of the future stream of payments. The 95% confidence interval for the implied discount rates estimates range between 2 and 6%, which are consistent with recent work on automobile purchases (e.g. Busse et al. (2013); Allcott and Wozny

(2014)). My findings suggest that since home buyers are attentive to and informed about fuel prices, pollution pricing policies such as taxes and cap-and-trade programs will create incentives not only to reduce the amount of energy people choose to consume, but to convert to cleaner heating fuels, and possibly increase the efficiency of building shells and appliances as well.

References

Allcott, H. and N. Wozny, "Gasoline Prices, Fuel Economy, and the Energy Paradox," *The Review of Economics and Statistics*, 2014, *96* (5), 779–795.

Alquist, R. and L. Kilian, "What Do We Learn from the Price of Crude Oil Futures?," *Journal of Applied Econometrics*, 2010, *25* (4), 539–573.

_ , _ , and R.J. Vigfusson, "Forecasting the Price of Oil," in G. Elliot and A. Timmerman, eds., *Handbook of Economic Forecasting vol. 2.*, Amsterdam: North-Holland: Elsevier, 2012.

Anderson, S., R. Kellogg, and J. Sallee, "What Do Consumers Believe About Future Gasoline Prices?," *Journal of Environmental Economics and Management*, 2013, *66*, 383–403.

Brounen, D. and N. Kok, "On the Economics of Energy Labels in the Housing Market," *Journal of Environmental Economics and Management*, 2011, *62*, 166–179.

Busse, M., C. Knittel, and F. Zettelmeyer, "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, 2013, *103* (1), 220–256.

Castaneda, C.J., "Regulated Enterprise: Natural Gas Pipelines and Northeastern Markets 1938-1954," The Ohio State University Press 1993.

Chetty, R., A. Looney, and K. Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

Davis, L. and L. Kilian, "he Allocative Cost of Price Ceilings in the U.S. Residential Market for Natural Gas," *Journal of Political Economy*, 2011, *119* (2), 212–241.

Dinan, T. and J. Miranowski, "Estimating the Implicit Price of Energy Efficiency Improvements in the Residential Housing Market: A Hedonic Approach," *Journal of Urban Economics*, 1989, *25*, 52–67.

Dubin, J. and M. McFadden, "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, 1984, *52* (2), 345–362.

Finkelstein, A., "E-ztax: Tax Salience and Tax Rates," *Quarterly Journal of Economics*, 2009, *124*, 969–1010.

Goldberg, P., "The Effects of the Corporate Average Fuel Economy Standards in the U.S.," *Journal of Industrial Economics*, 1998, *46*, 1–33.

Grigolon, L., M. Reynaert, and F. Verboven, "Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market," *Working Paper: University of Leuven*, 2014.

Hausman, J., "Individual Discount Rates and the Purchase and Utilization of Energy Using Durables," *Bell Journal of Economics*, 1979, *10* (1), 220–225.

Hossain, T. and J. Morgan, "Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay," *Advances in Economic Analysis and Policy*, 2006, *6* (2).

Houde, S., "How Do Consumers Respond to Product Certifications: A Welfare Analysis of the Energy Star Program," *Working Paper, University of Maryland*, 2014.

_ **and C.A. Spurlock**, "Do Energy Efficiency Standards Improve Quality? Evidence from a Revealed Preference Approach," *Working Paper: University of Maryland*, 2015.

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

Jessoe, K. and D. Rapson, "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use," *American Economic Review*, 2014, *104* (4).

Johnson, R. and D.L. Kaserman, "Housing Market Capitalization of Energy-Saving Durable Good Investments," *Economic Inquiry*, 1983, *21* (3), 374–386.

Myers, E., "Asymmetric Information in Residential Rental Markets: Implications for the Energy Efficiency Gap," *Energy Institute at Haas Working Paper*, 2017, (246).

Notte, J., "Demand Grows in N.E. for Natural Gas Heat," *The Boston Globe*, 25 April, 2012.

Rapson, D., "Durable Goods and Long-run Electricity Demand: Evidence from Air Conditioner Purchase Behavior," *Journal of Environmental Economics and Management*, 2014, *68* (1), 141–160.

Reiss, P. and M. White, "What Changes Energy Consumption? Prices and Public Pressure," *RAND Journal of Economics*, 2008, *39* (3), 636–663.

Sallee, J.M., S. West, and W. Fan, "Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations," *Journal of Public Economics*, 2016, *135*, 61–73.

Wooldridge, J., *Econometric Analysis of Cross Section and Panel Data, Second Edition*, MIT Press, 2010.

Tables

Table 1: Covariate Comparison Between Fuel Type

	Gas	Oil	P-value of Diff
Sale Price	$342,104	$322,718	0.00***
Number of Bedrooms	3.11	3.19	0.00***
Number of Bathrooms	2.36	2.20	0.00***
Number of Stories	1.78	1.73	0.00***
Square Feet	1912.90	1889.70	0.00***
Year Built	1956.59	1947.94	0.00***
Exterior Wall Type			
Wood	45%	46%	
Vinyl	32%	33%	
Aluminum	11%	12%	
Other	13%	9%	
Heat Type			
Forced Air	50%	26%	
Forced Hot Water	38%	60%	
Steam	8%	13%	
Other	3%	1%	
Observations	303,802	301,005	

Notes: Characteristic and transaction data are from CoreLogic for the state of Massachusetts (1990-2011). All prices are inflated to 2012 dollars.

Table 2: Estimation of the Effect of Relative Fuel Prices on Relative Transaction Prices

	Sales Price	Sales Price	Sales Price	Sales Price	Sales Price
Fuel Price	-1186.4***	-1002.4***	-1122.1***	-1074.7***	-1064.7***
	(198.8)	(242.6)	(115.1)	(167.1)	(131.7)
Oil Heat Indicator	-15334.4***	-8165.6***	1311.1		
	(1066.5)	(1157.8)	(978.3)		
Year FE	Yes	Yes	No	Yes	No
Attribute Controls	No	Yes	Yes	No	No
Geographic Area × Year FE	No	No	Yes	No	Yes
Unit FE	No	No	No	Yes	Yes
N	909434	870567	870504	529156	529008
Implied Discount Rate Infinite Horizon	8.4%	9.1%	8.6%	9.4%	9.5%

Notes: Regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic, years 1990-2011. Price is the average annual residential retail fuel price for oil or natural gas in dollars per MMBTU. All prices are inflated to 2012 dollars. Standard errors are two-way clustered at the house and geographic unit by year level and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.

Table 3: Estimation of the Effect of Relative Fuel Price on Relative Transaction Price: Additional Controls

	Sales Price	Sales Price
Fuel Price	-793.5***	-751.9***
	(184.8)	(165.0)
Unit FE	Yes	Yes
Geographic Area × Year FE	Yes	Yes
Oil Linear Trend	Yes	Yes
Age FE	No	Yes
N	529008	528642
Implied Discount Rate Infinite Horizon	13.1%	13.9%

Notes: Regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic years 1990-2011. Price is the average annual residential retail fuel price for oil or natural gas in dollars per MMBTU. All prices are inflated to 2012 dollars. Standard errors are two-way clustered at the house and geographic unit by year level and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.

Table 4: IV Estimation of the Effect of Relative Fuel Price on Relative Transaction Price

(Dependent Variable)	First Stage (Fuel Price)	2SLS (Sales Price)
Fuel Price IV	1.080***	
	(0.0271)	
Fuel Price		-3544.8***
		(1012.7)
F-stat	808.7	
R^2	0.886	
Unit FE	Yes	Yes
Geographic Area × Year FE	Yes	Yes
Oil Linear Trend	Yes	Yes
Age FE	Yes	Yes
N	528642	528642
Implied Discount Rate Infinite Time Horizon		2.7%

Notes: Regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic years 1990-2011. Price is the average annual residential retail fuel price for oil or natural gas in dollars per MMBTU. The instrument for price is the proportion of homes built with oil in the year a particular house was built times the oil price in the year it was sold plus the proportion of homes built with gas in the year a particular house was built times the gas price in the year it was sold. All prices are inflated to 2012 dollars. Standard errors are two-way clustered at the house and geographic unit by year level and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.

Table 5: Placebo IV Estimation of the Effect of Relative Fuel Price on Relative Transaction Price

(Dependent Variable)	First Stage (Placebo Fuel Price)	2SLS (Sales Price)
Fuel Price IV	3.766***	
	(0.0194)	
Placebo Fuel Price		-160.2
		(365.9)
F-stat	21697.5	
R^2	0.977	
Unit FE	Yes	Yes
Geographic Area × Year FE	Yes	Yes
Oil Linear Trend	Yes	Yes
Age FE	Yes	Yes
N	247932	247932

Notes: Only oil heated homes are included in this regression. A placebo oil indicator is defined as homes built in years where more than 50% of homes built are heated with oil. The placebo fuel price is defined as the oil price if the placebo oil indicator is 1 and the gas price otherwise. The instrument for placebo price is the proportion of homes built with oil in the year a particular house was built times the oil price in the year it was sold plus the proportion of homes built with gas in the year a particular house was built times the gas price in the year it was sold. All prices are inflated to 2012 dollars. Regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic years 1990-2011. Standard errors are two-way clustered at the house and geographic unit by year level and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.

Table 6: Estimation of the Effect of Relative Fuel Prices on Relative Transaction Prices

	Sales Price	Sales Price	Sales Price	Sales Price	Sales Price
Mean Futures Price	-1160.5***	-1028.1***	-1216.3***	-972.9***	-1058.6***
	(179.4)	(202.3)	(119.0)	(153.2)	(121.7)
Oil Heat Indicator	-15644.3***	-8458.9***	955.0		
	(1055.0)	(1174.8)	(1029.9)		
Year FE	Yes	Yes	No	Yes	No
Attribute Controls	No	Yes	Yes	No	No
Geographic Area × Year FE	No	No	Yes	No	Yes
Unit FE	No	No	No	Yes	Yes
N	909434	870567	870504	529156	529008
Implied Discount Rate Infinite Horizon	8.6%	9.8%	8.2%	10.4%	9.5%

Notes: Regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic, years 1990-2011. The Mean Futures Price is calculated by weighting all traded futures prices by the discount factor and is measured in dollars per MMBTU. All prices are inflated to 2012 dollars. Standard errors are two-way clustered at the house and geographic unit by year level and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.

Figures

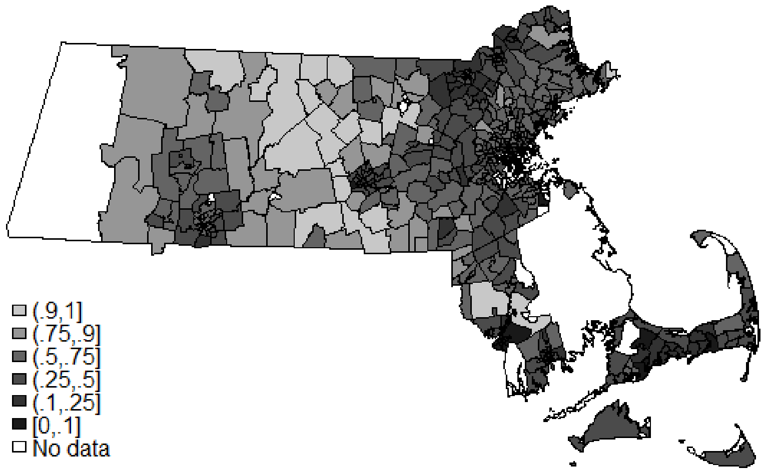

Figure 1: Proportion of Homes Heated With Oil

Notes: Housing characteristic data for the state of Massachusetts are from CoreLogic

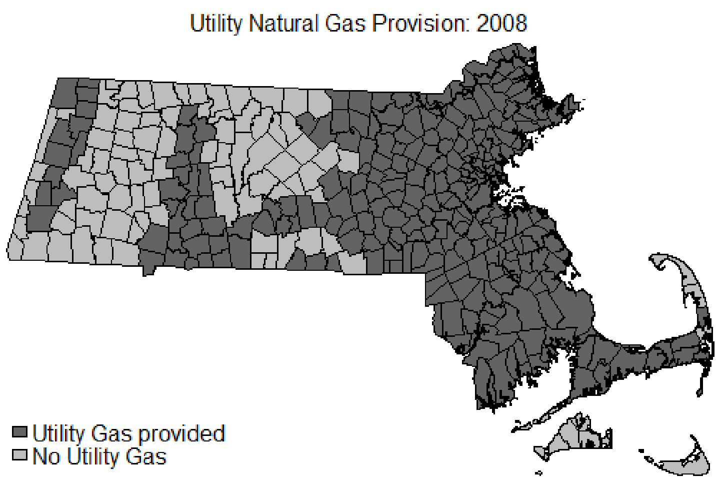

Figure 2: Utility Natural Gas Provision: 2008

Notes: Natural gas utility territory data for the state of Massachusetts are from MassGIS

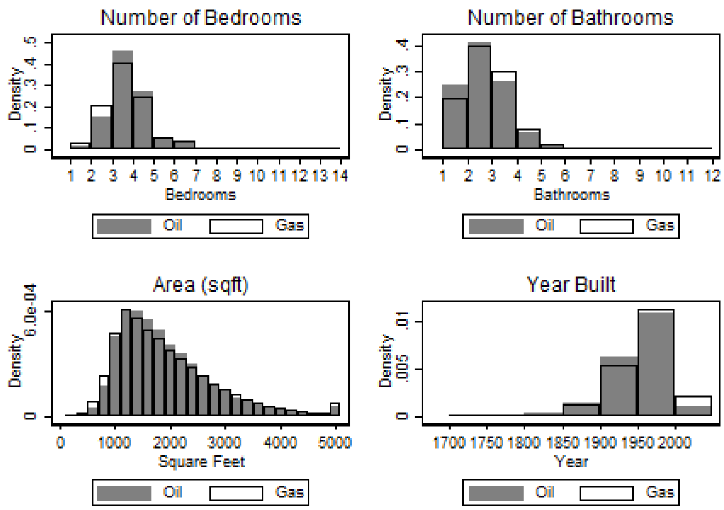

Figure 3: Overlap of Covariates

Notes: Housing characteristic data for the state of Massachusetts are from CoreLogic

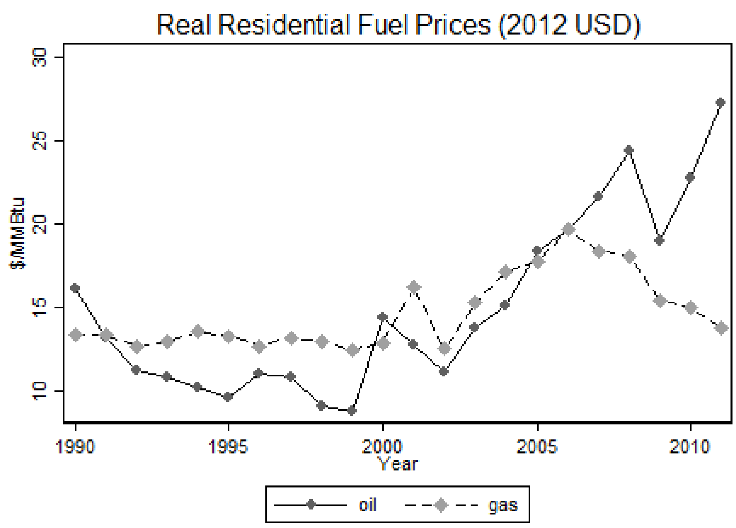

Figure 4: Real Residential Fuel Prices (2012 USD)

Notes: The prices are average annual retail prices ($/MMBTU) for the state of Massachusetts from EIA. All prices are inflated to 2012 dollars.

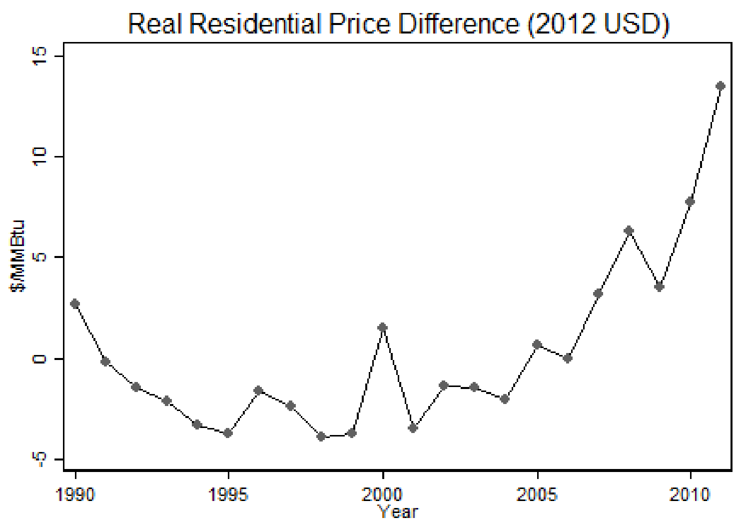

Figure 5: Real Residential Price Difference (2012 USD)

Notes: This figure displays the price of oil minus the price of natural gas. The prices are annual retail prices ($/MMBTU) for the state of Massachusetts from EIA. All prices are inflated to 2012 dollars.

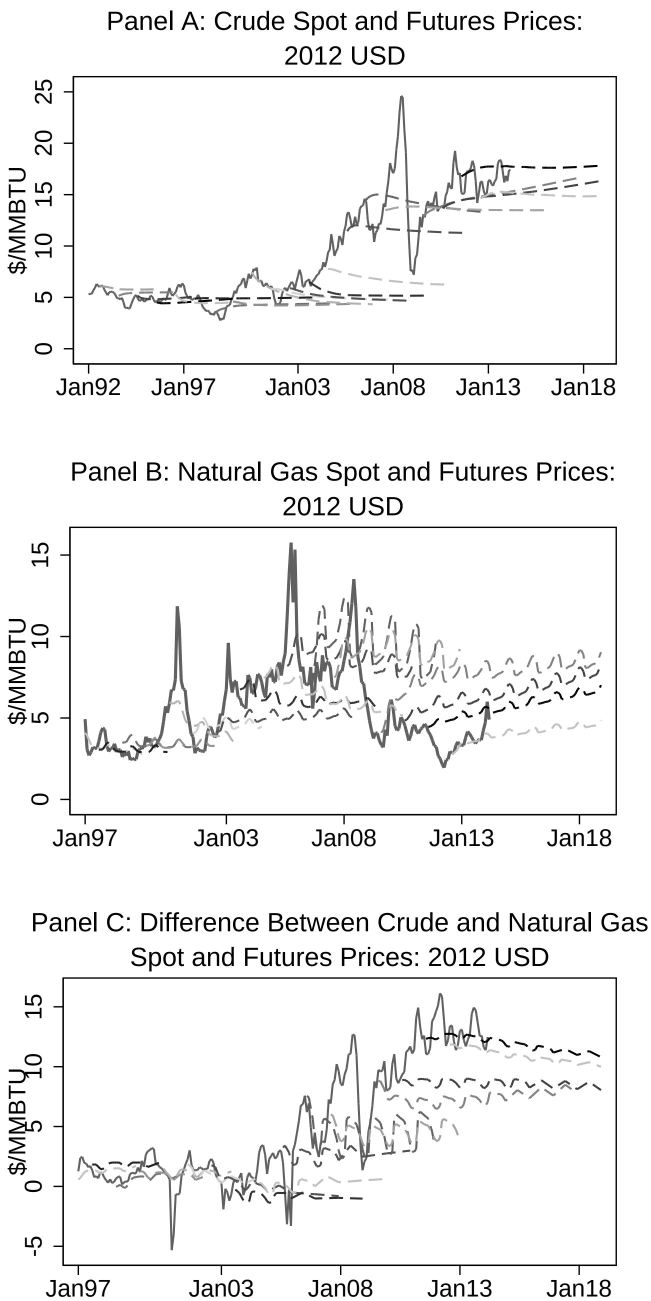

Figure 6: Spot and Futures Prices

Notes: The solid line in Panels A an B are the spot price and the dashed lines are forward curves taken every June. Panel C displays crude spot and futures prices minus natural gas spot and futures prices. All prices are in 2012 dollars. Forward curves are inflated according to the trade date. Source: NYMEX

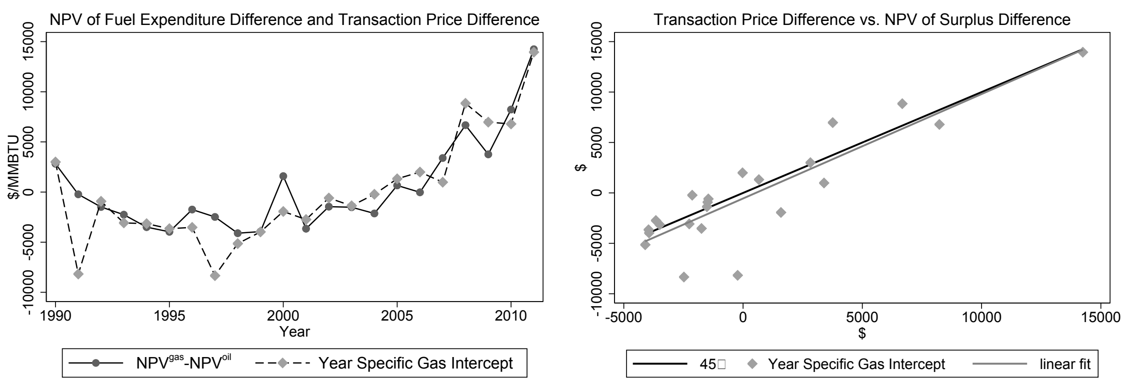

Figure 7: Net Present Value of the Fuel Expenditure Difference For Oil vs. Gas Houses Over Infinite Horizon With 9.5% Discount Rate and the Difference in Housing Transaction Prices

Notes: The graph on the left depicts the difference in the net present value of fuel expenditure between oil and gas houses and the difference in transaction prices. The graph on the right plots each mean difference in annual transaction price against the difference in the NPV of fuel expenditure between oil and gas houses for each year. All prices are inflated to 2012 dollars

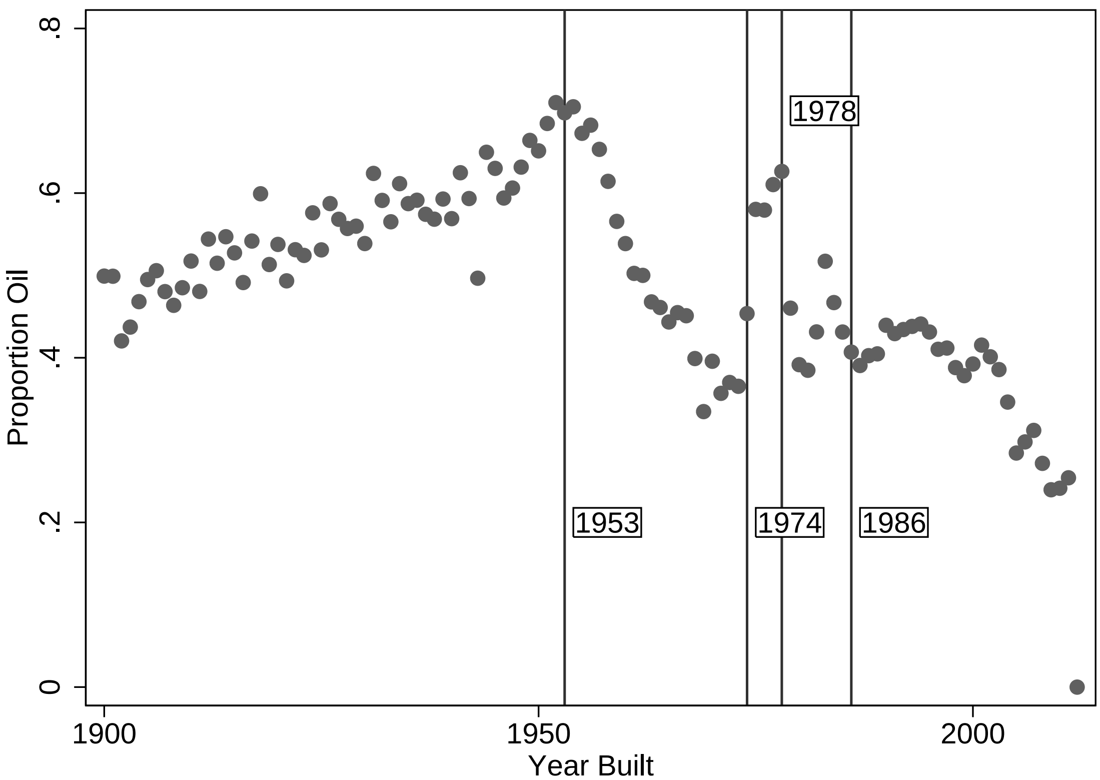

Figure 8: Proportion of Homes Built with Oil by Year Built

Notes: The graph depicts the proportion of homes of built with oil for each vintage year between 1900 and 2011. There are several clear breaks in the data, which are associated with policy changes that are exogenous to the local housing market. Starting in 1953 piped natural was imported into New England for the first time. Between 1974 and 1978, there was an increase in homes built with oil due to supply shortages of natural gas. In 1978, wellhead price controls for gas were lifted, which increased natural gas supply, and the proportion of homes built with oil dropped. Since then, natural gas became more common with the exception of a brief increase in homes built with oil in the mid-1980's following the crude oil price collapse of 1986. The housing transaction price data are provided by CoreLogic for the state of Massachusetts.

Appendix

For Online Publication

A1 Equivalence of Two Estimation Approaches

In what follows I describe the equivalence of two estimation approaches for estimating the difference-in-differences of interest: the difference in housing price between oil and gas heated homes from one year (e.g. high price difference) to the next (e.g. low price difference).

Let α be the individual effects of each fuel price/fuel type combination as follows:

$$\begin{aligned} Outcome_{jat} = \beta_0 + \alpha_1(\mathrm{I}^{\mathrm{oil}}_{jt} \times p^{\mathrm{oil}}_t) + \alpha_2(\mathrm{I}^{\mathrm{oil}}_{jt} \times p^{\mathrm{gas}}_t) + \alpha_3(\mathrm{I}^{\mathrm{gas}}_{jt} \times p^{\mathrm{oil}}_t) + \alpha_4(\mathrm{I}^{\mathrm{gas}}_{jt} \times p^{\mathrm{gas}}_t) \\ +\lambda_{at} + \theta_{ja} + \epsilon_{jat} \end{aligned} \tag{A1}$$

Following the notation in the paper *Outcome*, is the outcome for unit j in geographic area a in year t. The annual fuel price is p with oil and gas indicated with superscripts, "I" indicates the primary heating fuel with oil and gas as superscripts. λ_{at} are geographic area by year fixed effects, θ_{ja} are individual fixed effects, and ϵ_{jat} is the error term.

Factoring out the fuel indicators in Equation A1 yields the following estimation equation:

$$Outcome_{jat} = \beta_0 + \beta_1 \mathrm{I}^{\mathrm{oil}}_{jt} \times (p^{\mathrm{oil}}_t - p^{\mathrm{gas}}_t) + \beta_2 \mathrm{I}^{\mathrm{gas}}_{jt} \times (p^{\mathrm{oil}}_t - p^{\mathrm{gas}}_t) + \lambda_{at} + \theta_{ja} + \epsilon_{jat} \tag{A2}$$

The coefficient β_1 in the following equation estimates difference-in-differences of interest. Since one of these terms, e.g. $\mathrm{I}^{\mathrm{gas}}_{jt} \times (p^{\mathrm{oil}}_t - p^{\mathrm{gas}}_t)$ is collinear with year fixed effects, the estimate of $\beta_1 = (\alpha_1 - \alpha_2) - (\alpha_4 - \alpha_3)$.

The estimation procedure used in the paper yields equivalent results, where this time the terms from Equation A1 are gathered slightly differently:

$$Outcome_{jat} = \beta_0 + \beta_1(\mathrm{I}_{jt}^{\text{oil}} \times p^{oil} + \mathrm{I}_{jt}^{\text{gas}} \times p^{gas}) + \beta_2(\mathrm{I}_{jt}^{\text{gas}} \times p^{oil} + \mathrm{I}_{jt}^{\text{oil}} \times p^{gas}) + \lambda_{at} + \theta_{ja} + \epsilon_{jat} \tag{A3}$$

Again, one term will be collinear with year fixed effects, e.g. $\mathrm{I}_{jt}^{\text{gas}} \times p^{oil} + \mathrm{I}_{jt}^{\text{oil}} \times p^{gas}$. I simplify the first term to p_t since $p_t = \mathrm{I}_{jt}^{\text{oil}} \times p^{oil} + \mathrm{I}_{jt}^{\text{gas}} \times p^{gas}$. β_1 yields and equivalent difference in differences, i.e. $\beta_1 = (\alpha_1 - \alpha_4) - (\alpha_3 - \alpha_2)$.

A2 Description of Energy Expenditure Data

In what follows I describe the Energy Expenditure Data used for the two-stage procedure detailed in the next section (A3). RECS is an in-home survey, which provides detailed information on housing unit characteristics as well as energy usage and expenditures by fuel type and end-use. The price and expenditure data are verified with households' residential energy suppliers to ensure their reliability.

The survey is conducted approximately every five years and is designed to be a nationally representative cross-section of U.S. housing units. I use data from 6 surveys performed between 1990 and 2009 in my analysis. I use data from the Northeast Census region to predict energy expenditure as a function of household income controlling for size and other housing characteristics. I limit the sample to owner-occupied, single family houses in the northeast census region–a total of 1545 housing units.

A3 Two Stage Approach

I estimate a two sample two-staged least squares (TS2SLS) model, a variant of the two sample instrumental variables (TSIV) procedure discussed in Angrist and Krueger (1992, 1995).[17] This two-sample IV procedure addresses concerns about measurement error in the NPV of fuel expenditure, but the limitation still exists that the only exogenous covariates that can be included as controls must be present in both samples.

In the first stage, I use the RECS data to estimate the effect of heating fuel price movements on energy expenditures as a function of unit characteristics. Then, I estimate a variant of Equation 3 by regressing housing price from the CoreLogic data on the estimated expenditure and unit characteristics. The first stage of the estimation is as follows.

First Stage

$$Exp_{jt} = \beta_0 + \beta_1 p_t + \beta_2 \mathrm{I}^{\text{oil}}_{jt} + \mathbf{X}_{jt}\beta + \delta_t + \epsilon_{jt} \tag{A4}$$

The dependent variable is expenditure on the primary heating fuel, Exp, for unit j in survey year t. The annual fuel price is p, $\mathrm{I}^{\text{oil}}_{it}$ indicates oil as the primary heating fuel. As with the main estimation, I use one statewide average price for each of these fuels. δ_t are year fixed effects. X_{jt} is a matrix of covariates and ϵ_{jt} is the error term. The covariates for this estimation include, the number of rooms, bathrooms and stories, flexible controls for square footage, binned by 1000 square foot increments, and indicators for decade built. They were chosen because they were available and in both surveys and are comparable between the two samples. The second stage estimation is as follows, where estimates of expenditure, $\hat{Exp}_{jt}$, are a function of the house characteristics in the transactions data and the coefficients estimated in the first stage. The coefficient of interest, $\hat{\gamma}$ can be interpreted as the effect of a \$1 increase in the present value of annual fuel expenditure on housing price.

[17] Angrist and Krueger (1992) show that consistent instrumental variables estimation is still possible if one sample contains the outcome, another distinct sample contains the exogenous regressor, and both samples contain the instrumental variable and other exogenous variables included in the model. Their two sample instrumental variables (TSIV) estimator is: $\hat{\beta}_{TSIV} = (Z_2' X_2/n_2)^{-1}(Z_1' Y_1/n_1)$, where Y is the outcome, X contains the endogenous regressor and other exogenous variables, and Z is the matrix of valid instrumental variables, n is the number of observations and subscripts denote the samples 1 and 2. Inouue and Solon (2010) show that $\hat{\beta}_{TS2SLS} = (\hat{X}_1'\hat{X}_1)^{-1}\hat{X}_1' Y_1$ and $\hat{\beta}_{TSIV}$ as proposed by Angrist and Krueger (1992, 1995) have the same probability limit, though TS2SLS is more asymptotically efficient in finite samples due an implicit correction for differences in the distribution of Z between the two samples.

Second Stage

$$H_{jt} = \alpha_0 + \hat{\gamma}\hat{Exp}_{jt} + \alpha_1 \mathrm{I}^{\mathrm{oil}}_{jt} + \mathbf{X}_{jt}\beta + \delta_t + \epsilon_{jt} \tag{A5}$$

The results from the estimation are displayed in Table A1. The first stage estimate shows that a \$1 increase in the annual MA residential fuel price leads to an increase of \$100 in annual expenditure for houses in the northeast census region. The reduced form estimate and the implied discount rate of 10% are quite close to the basic estimation in the main analysis.

A4 Estimation Using Futures Prices

I examine the sensitivity of the basic estimation to using the discount factor-weighted mean of future fuel prices rather than contemporaneous price. Crude oil and natural gas are traded for as much as 7 to 13 years in advance for later years in the sample. These years will have the largest impact on perceived prices in NPV terms. The discount factor-weighted mean is constructed as follows.

$$\frac{\sum_{i=t}^{T} \delta^i \cdot p_{ijt}}{\sum_{i=t}^{T} \delta^i} \tag{A6}$$

Future periods are indexed by i, the discount factor is δ^i, and the fuel price for house j for future year i in year t is p_{ijt}. If the time horizon of the decision were limited to the number of future periods that derivatives are traded T, this would be the perceived price for an agent using the futures market to forecast price. I inflate forward prices according to the trade date and transform crude oil and natural gas prices into residential heating oil and natural gas prices using the average historical relationships between the traded fuel and the residential price. Specifically, I predict the average historical relationship using simple linear regressions of levels of residential retail prices on levels of crude oil or Henry Hub natural gas spot prices. The reason I do this is residential gas prices do not have a separate futures market and residential heating oil is not traded for time horizons of more than 2 to 3 years.

A sensitivity test using the discount factor-weighted average futures prices incorporates all of the information available in the futures market, but implicitly assumes that home buyers will use the discount-weighted average of the prices in traded years for periods beyond the last year for which there are trade data. I replicate Table 2 using the discount factor-weighted average futures price. The results are displayed in Table A2. The magnitude of the point estimates are quite close to those using contemporaneous prices. For the preferred specification (column 5) with geographic area by time and house fixed effects, the estimates using the discount factor-weighted mean futures price actually yield the same implied discount rate of 9.5% as the estimates using contemporaneous price. Futures prices tend to not deviate too far from spot prices, meaning even if consumers were paying attention to them, their decisions would not deviate significantly from a consumer using contemporaneous prices.

Appendix Tables

Table A1: Two Sample 2SLS: Estimation of the Effect of Relative Annual Energy Costs on Relative Transaction Prices Instrumented with Fuel Price

	First Stage	Reduced Form	Second Stage
(Dependent Variable)	(Annual Fuel Expenditure)	(Sales Price)	(Sales Price)
Fuel Price	100.4***	-1082.4***	
	(12.92)	(291.0)	
Estimated Annual Fuel Expenditure			-10.79**
			(4.89)
Oil Heat Indicator	-157.6***	-6067.9	-7767.7
	(39.23)	(4852.7)	(5144.28)
F-stat	20.42		
R^2	0.363		
Attribute Controls	Yes	Yes	Yes
N	1515	909434	909434
Implied Discount Rate Infinite Horizon			10%

Notes: First stage regression data are from the Residential Energy Consumption Survey (RECS), northeast census division, survey years 1990, 1993, 1997, 2001, 2005, and 2009. the second stage regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic, years 1990-2011. Price is the average annual residential retail fuel price for oil or natural gas in dollars per MMBTU. All specifications control flexibly for the house vintage, number of rooms, bedrooms and bathrooms, square footage and year fixed effects. All prices are inflated to 2012 dollars. Standard errors are bootstrapped with 10,000 iterations, clustered at geographic area, and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.

Table A2: Estimation of the Effect of Relative Fuel Prices on Relative Transaction Prices

	Sales Price	Sales Price	Sales Price	Sales Price	Sales Price
Mean Futures Price	-1160.5***	-1028.1***	-1216.3***	-972.9***	-1058.6***
	(179.4)	(202.3)	(119.0)	(153.2)	(121.7)
Oil Heat Indicator	-15644.3***	-8458.9***	955.0		
	(1055.0)	(1174.8)	(1029.9)		
Year FE	Yes	Yes	No	Yes	No
Attribute Controls	No	Yes	Yes	No	No
Geographic Area × Year FE	No	No	Yes	No	Yes
Unit FE	No	No	No	Yes	Yes
N	909434	870567	870504	529156	529008
Implied Discount Rate Infinite Horizon	8.6%	9.8%	8.2%	10.4%	9.5%

Notes: Regressions are based on transaction and unit characteristic data for the state of Massachusetts from CoreLogic, years 1990-2011. The Mean Futures Price is calculated by weighting all traded futures prices by the discount factor and is measured in dollars per MMBTU. All prices are inflated to 2012 dollars. Standard errors are two-way clustered at the house and geographic unit by year level and are in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels.