The Mortality Effects of Winter Heating Prices

By

Janjala Chirakijja, Monash University
Seema Jayachandran, Princeton University and NBER
Pinchuan Ong, National University of Singapore Business School

Griswold Center for Economic Policy Studies
Working Paper No. 305, January 2023

The Mortality Effects of Winter Heating Prices*

Janjala Chirakijja[1], Seema Jayachandran[2,3], and Pinchuan Ong[4]

[1]Monash University, Department of Econometrics and Business Statistics
[2]Princeton University, Department of Economics & School of Public and International Affairs
[3]National Bureau of Economic Research
[4]National University of Singapore Business School

January 1, 2023

Abstract

This paper examines how the price of home heating affects mortality in the US. Exposure to cold is one reason that mortality peaks in winter, and a higher heating price increases exposure to cold by reducing heating use. Our empirical approach combines spatial variation in the energy source used for home heating and temporal variation in the national prices of natural gas and electricity. We find that a lower heating price reduces winter mortality, driven mostly by cardiovascular and respiratory causes. Our estimates imply that the 42% drop in the natural gas price in the late 2000s, mostly driven by the shale gas boom, averted 12,500 deaths per year in the US. The effect appears to be especially large in high-poverty communities.

*Chirakijja: janjala.chirakijja@monash.edu; Jayachandran: jayachandran@princeton.edu; Ong: ong.pinchuan@nus.edu.sg. We thank Sachet Bangia, Jamie Daubenspeck, Jacob Gosselin, Alejandro Favela, and Caitlin Rowe for outstanding research assistance, and several seminar participants and anonymous referees for helpful comments. Research reported in this paper was supported by the National Institute on Aging of the National Institutes of Health under award number R03AG058113. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

1 Introduction

Many families worldwide struggle to heat their homes each winter. Their heating bills are so high relative to their income that they are considered to be living in "fuel poverty." In the European Union, 8% of households are unable to keep their homes adequately warm in winter (Eurostat 2021). In the United States, 17% of households spend over 10% of their income on home energy; winter heating is the largest contributor (RECS 2009). The problem becomes even more acute during energy crises. For example, when natural gas supply was disrupted after Russia's invasion of Ukraine in 2022, heating prices soared in many parts of the world, pushing millions of additional households into fuel poverty.

Households face a difficult trade-off when heating prices are high: They have to keep their home uncomfortably cold to save on heating, or they have to forgo other spending to afford their high heating bill. Either choice could be harmful to their health. Using less heating means exposure to lower ambient temperature, which has been linked to cardiovascular, respiratory, and other health problems. But if families do not cut back heating usage one-for-one when the price rises, their energy bills will increase, leaving less money for other expenditures that affect health such as food and health care. For these reasons, morbidity and mortality are potentially important consequences of high heating prices.

This paper estimates the effect of heating prices on mortality in the US. A large literature has documented that mortality peaks in winter (see Appendix Figure A1) and that cold weather is associated with higher mortality. Our contribution is to examine whether high home heating costs exacerbate this pattern of "excess winter mortality."

Our empirical design uses spatial variation across the US in the energy source used for home heating. Natural gas and electricity are used for heating by 58% and 30% of US households, respectively. Importantly, there is considerable variation across counties in whether natural gas versus electricity is mainly used. We combine this spatial variation with temporal variation in the national prices of natural gas and electricity. The price of natural gas varied substantially over the 2000 to 2010 study period, relative to the price of electricity; it first rose, partly due to supply disruption from Gulf of Mexico hurricanes, and then fell after 2005, mostly due to the supply influx from shale production of natural gas

(Hausman and Kellogg 2015). We use the fact that when the price of natural gas rose or fell, households in areas that rely on natural gas for heating experienced a rise or fall in their home heating price, relative to households in areas reliant on electricity.

We find that lower heating prices reduce mortality in winter months.[1] The estimated effect size implies that the 42% drop in the price of natural gas in the late 2000s averted 12,500 winter deaths per year in the US. Moreover, we find that this effect does not just represent a short-run postponement of mortality. We also show that the effect, which is driven mostly by cardiovascular and respiratory causes and is larger in high-poverty communities, is robust to several stress tests of the empirical specification.

Our findings have implications for policies that reduce households' heating costs such as the federal Low Income Home Energy Assistance Program (LIHEAP) and state energy price subsidy programs in the US (see, e.g., Hahn and Metcalfe (2021)) and analogous policies worldwide, and are also relevant for cost-benefit analysis of weatherization programs that reduce households' need for heating. In addition, our findings highlight a health benefit of increases in US energy supply that has not received much prior attention.

Our paper contributes to the literature on the effects of cold weather on mortality (Eurowinter Group 1997; Analitis et al. 2008; Deschênes and Moretti 2009) and other dimensions of well-being (Ye et al. 2012; Bhattacharya et al. 2003; Cullen et al. 2004; Beatty et al. 2014). To our knowledge, no prior study has estimated the causal effect of heating prices — an important and policy-relevant mediating factor — on health. Previous work has found that the winter spike in mortality is especially large for people living in older housing, which tends to be poorly insulated, which is suggestive but not dispositive that indoor temperature is a mediating factor (Wilkinson et al. 2007).

Another related line of research examines adaptations that mitigate the temperature-health relationship. Previous research has examined the role of technological and medical advances (Barreca et al. 2016; Deschênes and Greenstone 2011), migration (Deschênes and Moretti 2009), and weatherization and energy-efficiency programs (Critchley et al. 2007; El Ansari and El-Silimy 2008; Green and Gilbertson 2008; Howden-Chapman et al. 2007).

[1] We define "winter" as November to March, the coldest months of the year in the US (see Appendix Figure A1).We also show the results using December to March, similar to analyses of excess winter mortality in the UK and Europe where those are the coldest months (Wilkinson et al. 2004)

Increased heating use is another important household-level adaptation, and we contribute by analyzing how high fuel prices stymie this adaptation. A study concurrent to ours analyzes the aftermath of the Fukushima nuclear power plant accident in Japan and finds that higher electricity prices exacerbate the relationship between cold temperatures and mortality (Neidell et al. 2021). An advantage of our research design is that we can directly identify changes in the price of heating (by incorporating geographic variation in the energy source used for heating) instead of energy prices more broadly, which might also affect health through other channels. Additionally, we shed light on the relative importance of the different mechanisms through which a higher heating price increases mortality.[2]

Our paper also contributes to the literature on the health effects of the shale gas (or "fracking") boom by highlighting a national-level health benefit — the drop in energy prices reduced winter mortality. Prior work has highlighted the health benefit of fracking displacing pollutive coal in electricity generation (Cullen and Mansur 2017; Fell and Kaffine 2018; Holladay and LaRiviere 2017; Knittel et al. 2015; Linn and Muehlenbachs 2018). Fracking has also been shown to be harmful because of local contamination from the chemicals used (Jackson et al. 2014; Groundwater Protection Council 2009; Muehlenbachs et al. 2015; Casey et al. 2016; Currie et al. 2017; Hill 2018). The health harm from the toxic chemicals is likely much larger per person affected than the health benefits from lower energy prices; however, the latter channel affects a much larger population. Thus, the net health effect of fracking aggregated for the whole US population is ambiguous. Finally, our empirical strategy is similar to that of Myers (2019) who compares households that use heating oil or natural gas in Massachusetts to study whether home energy costs are capitalized into home values.

2 Empirical strategy

To identify the effect of heating prices on mortality, we combine information on whether a locality typically uses natural gas or electricity for heating with data on national energy prices. This approach enables us to control for average differences across localities and time.

[2]Other studies have focused on financial assistance for energy bills or heating subsidies for low-income families (Frank et al. 2006; Grey et al. 2017; Crossley and Zilio 2018).

2.1 Estimating equations

In principle, we want to estimate the following equation:

$$\log(m_{jt}) = \alpha + \beta \log(p^H_{jt}) + \epsilon_{jt}. \tag{1}$$

Each observation is a county-month. The outcome $\log(m_{jt})$ is the log of age-adjusted mortality in county j in month t. (We use the log of the mortality rate following Stevens et al. (2015), but also report the results in levels.) The key regressor is $\log(p^H_{jt})$, the log of the heating price for the county-month. The coefficient β measures the elasticity of mortality with respect to the heating price. The hypothesis is that $\beta > 0$: A higher heating price increases mortality.

There are no data on p^H_{jt} because utilities do not set a price specifically for heating, just for different energy sources. Instead, we construct a proxy for the heating price by interacting ShareGas_{jt}, the proportion of households in the area that used natural gas for heating in that year, with $\log(\text{RelPrice}_{jt})$, the ratio of the price of gas to electricity in the state-month. To see why this interacted variable tracks the heating price for households, note that when natural gas prices increase (high RelPrice), areas with high ShareGas face relatively higher heating prices. Conversely, when electricity prices increase (low RelPrice), areas with higher ShareGas face relatively low heating prices. In practice, most of the identifying variation comes from the natural gas price because it fluctuates more over the study period.[3]

Utilities markets within the US vary considerably in terms of prices and regulations, which means that $\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt})$ could be endogenous to local demand. To solve this problem, our empirical strategy relies on national-level energy prices combined with (pre-period) local variation in the energy source for heating. That is, we instrument for $\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt})$ with $\text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t})$.[4]

[3]Our results are similar if we replace RelPrice with the price of natural gas, with or without controlling for ShareGas interacted with the electricity price.

[4]Formally, $\text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t}) = \text{ShareGas}_{j,2000} \log(p^G_{\text{US},t}) + \text{ShareElec}_{j,2000} \log(p^E_{\text{US},t}) - \log(P^E_{\text{US},t})$, where $\text{ShareElec}_{j,2000}$ is the proportion of households in 2000 that use electricity for heating, and $p^G_{\text{US},t}$ and $p^E_{\text{US},t}$ are the national prices of natural gas and electricity, respectively. Month-year fixed effects absorb $\log(p^E_{\text{US},t})$. The first two terms on the right capture the average proportional change in the heating price *across* households in a county (some uses gas, while others use electricity as their main heating source), i.e., it is an exogenous proxy for $\log(p^H_{jt})$.

We estimate the following equation with this instrumental variables approach:

$$\log(m_{jt}) = \alpha + \beta \,\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt}) + \gamma_j + \tau_t + \theta Z_j \times \log(\text{RelPrice}_{\text{US},t}) \\ + \delta X_{jt} + \epsilon_{jt}, \tag{2}$$

where the first stage equation is as follows:

$$\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt}) = \tilde{\alpha} + \tilde{\beta}\, \text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t}) + \tilde{\gamma}_j + \tilde{\tau}_t \\ + \tilde{\theta} Z_j \times \log(\text{RelPrice}_{\text{US},t}) + \tilde{\delta} X_{jt} + \nu_{jt}. \tag{3}$$

In addition to replacing $\log(p^H_{jt})$ with $\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt})$, we augment equation (1) by including county fixed effects, γ, and month-year fixed effects, τ. We also include several control variables, denoted by the vector X. Because the study period spans the housing market boom and bust as well as the Great Recession, we control for a housing price index, the unemployment rate, and the manufacturing share of local employment income. X also includes factors that might affect mortality, namely air pollution — particulate matter 2.5 and 10 microns, separately, and nitrogen dioxide — absolute humidity, and the heating degree-days (HDD) of the area (a measure of coldness, described in Section 3). We additionally include nitrogen dioxide as a quadratic term to control for it more flexibly because we find that it is correlated with $\text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t})$. The humidity-mortality relationship is non-linear (Barreca and Shimshack 2012), so we also control for a quadratic term in absolute humidity. Finally, we control for area characteristics Z, specifically pre-period log income (25th, 50th, and 75th percentiles) and the share of the population over age 70, interacted with $\log(\text{RelPrice}_{\text{US},t})$; these controls help safeguard against a spurious correlation due to the Great Recession (or another phenomenon with a similar temporal pattern as $\log(\text{RelPrice}_{\text{US},t})$) having a differential impact on mortality across socioeconomic or demographic groups (Hoynes et al. 2012).

The identification assumption is that when natural gas prices are high relative to electricity, places with more natural gas usage for heating have higher mortality only because of the higher heating price they face, conditional on fixed effects and control variables.

Throughout, we cluster standard errors by state to allow for serial correlation plus spatial correlation among counties in a state.

For our baseline specification, we restrict the data to only winter months (when possible) when most of the year's heating is consumed. We also estimate a winter/non-winter specification that uses the non-winter months as an additional comparison group, testing the prediction that the price of heating affects mortality more in winter than in the remaining, warmer months:

$$\begin{aligned} log(m_{jt}) \quad = \quad & \alpha + \lambda_1 \, \text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt}) \times \text{Winter}_t \\ & + \lambda_2 \, \text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt}) \\ & + \lambda_3 \, \text{ShareGas}_{j,2000} \times \text{Winter}_t + \lambda_4 \log(\text{RelPrice}_{\text{US},t}) \times \text{Winter}_t \\ & + \theta_1 Z_j \times \log(\text{RelPrice}_{\text{US},t}) \times \text{Winter}_t + \theta_2 Z_j \times \log(\text{RelPrice}_{\text{US},t}) \\ & + \theta_3 Z_j \times \text{Winter}_t + \gamma_j + \tau_t + \delta X_{jt} + \epsilon_{jt} \end{aligned} \tag{4}$$

Analogous to before, the first two regressors are instrumented using $\text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t})$ and $\text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t}) \times \text{Winter}_t$. The prediction is $\lambda_1 > 0$.

Some winters or particular months in winter are colder than others, so we can also replace Winter with HDD. In this specification, we control for the county's average HDD in winter, $\overline{\text{HDD}}_j$, in parallel to HDD_{jt} to adjust for systematic differences (e.g., demographics) between colder regions such as the Midwest and warmer ones such as the South.

2.2 Assessing the heating and non-heating consumption channels

Heating prices can affect mortality through two channels: a cutback in heating use ("heating channel") and a reduction in the income left over for other consumption after paying the heating bill ("non-heating channel"). To gauge the potential relevance of each channel, we analyze two additional outcomes.

The first one is the (log) quantity of home energy use. Here, the coefficient β from equation (2) can be interpreted as a price elasticity. We expect it to be negative: Consumers substitute away from heating when it becomes more expensive. The data on home energy use do not disaggregate it by purpose (e.g., heating, lighting). Thus, while the variation in the price of natural gas is mainly measuring variation in a household's heating price, the outcome combines heating plus other energy uses, so the coefficient represents a lower bound magnitude for the price elasticity of heating demand. The use of natural gas in homes is mostly for heating (space heating and water heating), with an additional small contribution from kitchen ranges and clothes dryers. Non-heating home energy needs such as lighting, refrigeration, and air conditioning predominantly use electricity throughout the US. Home heating is the largest component of home energy use, accounting for 42% of annual home energy consumption, with water heating accounting for an additional 18% (RECS 2009).

The second outcome is expenditures on home energy, again with the caveat that we cannot distinguish spending on heating from other energy uses (although in winter months, heating accounts for most energy use). If households are not cutting back one-for-one when the price rises, then higher energy prices will lead to higher energy bills (and thus less income left for other consumption).

2.3 Geographic variation in heating source

Natural gas and electricity are the two most common energy sources for home heating in the US, with considerable geographic variation. In some communities, almost every household uses natural gas for heating, and in other communities, almost no one does.

Figure 1 shows the share of households using natural gas as their heating source across counties, based on 2000 US Census data.

Whether a locality uses natural gas, electricity, or another heating source is not random, and various factors explain the differences. Natural gas pipelines do not extend to some parts of the US, such as Maine. Areas that are well-suited for hydroelectric power generation have low electricity costs and thus rely more on electricity. For historical reasons, much of the Northeast uses heating oil, a petroleum product, instead of gas or electricity. Importantly, the geographic differences were determined long before the study period and are highly persistent. Being predetermined does not rule out that an area's heating source is correlated with other factors affecting mortality, so the analysis controls for other locality characteristics in parallel to heating source. This guards against the endogeneity of shares emphasized by

Goldsmith-Pinkham et al. (2020).[5]

2.4 Temporal variation in energy prices

Figure 2 plots the national prices of natural gas and electricity over the 2000 to 2010 study period. The data source is the US Energy Information Administration (EIA). Natural gas is one of the fuel sources used in electricity generation, so the two prices co-move, but far from in lockstep. Electricity prices changed somewhat over the time period, while natural gas prices rose and then fell much more dramatically. As a result, the relative price of natural gas to electricity rose and then fell over the period.

Natural gas prices rose from 2004 to 2005 due in part to supply disruptions from major hurricanes along the Gulf coast (Hurricane Ivan in 2004 and Hurricanes Katrina and Rita in 2005) (Brown and Yücel 2008). In addition, increased efficiency of producing electricity from natural gas boosted demand for natural gas during the early 2000s (Hartley et al. 2008). A main cause of the natural gas price decline in the mid-2000s was the sharp increase in shale gas production (plotted in Figure 2); Hausman and Kellogg (2015) estimate that increased supply from shale gas explains 83% of the 2007-2013 decline in the price of natural gas.[6]

2.5 Home heating versus other heating

While we sometimes refer to our results as due to home heating, the analysis cannot isolate home heating from other indoor (e.g., workplace) heating. Some policy implications, such as whether to promote increased energy supply, are similar whether the channel is home heating or other indoor heating. For other policies, such as subsidies for consumer heating bills, it would be valuable to isolate heating costs at home, which our research design does not permit. A related, more minor limitation is that we cannot separate the effect of space heating from water heating; the energy source is the same in most households (RECS 2014), and both types of heating likely affect health through similar mechanisms.

[5] Users of natural gas can partially substitute to electric space heaters in the short run, but there is no low-cost short-run way to substitute in the other direction. In Appendix Table A1, we find little evidence of changes in heating source in response to changes in relative prices.

[6] To investigate whether the price decline is also due to lower demand for natural gas during the Great Recession, we estimated the relationship between RelPrice_{jt} and the unemployment rate (a proxy for the Great Recession intensity). The regression coefficient is small and statistically insignificant (see Appendix Table A2).

3 Data

Our analysis focuses on the contiguous US between 2000 and 2010. This section describes our data sources, with further details in Appendix B.

3.1 Mortality

We construct the county-year-month age-adjusted mortality rate from restricted-use Vital Statistics microdata. We exclude counties with a small population over age 50, specifically those in the bottom decile of counties, as they have few (often zero) deaths per month.[7]

We focus on causes of mortality that exhibit a high degree of excess winter mortality (EWM). Overall mortality is higher in winter than the rest of the year, but the pattern is more pronounced for some causes than others. We zero in on these causes because it is most plausible that they are exacerbated by exposure to cold and also because doing so increases statistical power. We use a data-driven approach to determine these causes. Using monthly data, we estimate a regression of log age-adjusted mortality for the entire US on a dummy for winter, separately for each of the 113 National Center for Health Statistics (NCHS) Selected Causes of Death. Causes with a large positive winter coefficient have more excess mortality in winter. We also estimate the model in levels to exclude minor causes that might have spuriously large coefficients. We select the causes whose winter coefficients are in the top quartile in both levels and logs, excluding two causes where there is no clear direct physiological link to cold exposure ("deaths from smoke, fire, and flames" and the residual category, "all other diseases"). The final 14 causes are within four alphabetic (i.e., broad) categories, and generally match the causes highlighted in the literature as exacerbated by cold (e.g., cardiovascular, respiratory). These high-EWM causes (hereafter, EWM causes) account for 61% of total mortality and 63% of total mortality in winter. Appendix Table A3 lists the 14 EWM causes, and Appendix Figure A2 shows the seasonality for EWM and non-EWM causes.

[7] These small counties constitute 0.37% of total population and 0.45% of total deaths in 2000. Among our retained counties, less than 0.03% of all county-month observations have zero deaths.

3.2 Independent variables

We construct county-level $\text{ShareGas}_{j,2000}$ using the 2000 Census Summary Files. For subsequent years, we use the American Community Survey (ACS), which is available starting in 2005, and linearly interpolate for years without data. (ShareGas is highly correlated over time — the correlation between ShareGas in 2000 and 2010 is 0.95.)

RelPrice, the ratio of the price of gas to electricity, is constructed using monthly state (for the endogenous heating price proxy) and national (for the instrument) price data from EIA. The appropriate specification depends on the timing of consumers' response to RelPrice. Similar to Auffhammer and Rubin (2018), we find that residential energy use responds to RelPrice with a lag of three months. Consumers seem to cut back on usage only after seeing their energy bill, which typically arrives a few weeks after the billing period ends. In addition, the health effects of cutbacks in heating use or paying higher bills might not be instantaneous. Hence, we use the average of the three- and four-month lagged price to construct RelPrice. We find similar results when we reduce the lag by one month or use annual prices. To investigate if the mortality effects materialize with a longer delay, we also estimate models that incorporate mortality effects in subsequent, post-winter months; the effect in subsequent months could also be negative if deaths are hastened by only a short duration ("harvesting").

The analysis also incorporates temperature data. We use daily average temperature (PRISM Climate Group 2004) to compute the heating degree-days (HDD) for each county-month. HDD is a commonly used measure of coldness — or need for heating — based on the idea that heating demand is linear in temperature when the temperature falls below 65°F. That is, $\text{HDD}_{jt} = \sum_{x=1}^{T} \max\{65 - tmean_{jtx}, 0\}$ where $tmean$ is the mean temperature of area j on day x of month t, and T is the number of days in month t. Appendix B provides details on the data sources for our control variables.

3.3 Other dependent variables

An auxiliary outcome we examine is the average price of home energy that consumers face. Our specification uses $\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt})$ as a proxy for the home *heating* price faced by households. We do not have household-level data on heating prices, but we can use aggregate administrative data on residential energy prices to verify that our regressor is a good proxy for household heating prices. The dependent variable we use for this is the weighted average of the local prices of natural gas and electricity, where weights are the local consumption levels of each energy source. Price and usage data are aggregated state-month-level data from EIA.

As discussed in Section 2.2, we also examine residential energy use. We sum natural gas and electricity usage from EIA data. To examine household spending on home energy, we combine 2000 Census microdata and ACS data for 2005 to 2010, aggregated to the county-year level.

4 Results

We first present results on the intermediate outcomes of home energy prices, quantity of energy consumed, and energy bills. We then present the mortality results.

4.1 Effect of heating price on energy use and spending

We start by examining the usage-weighted average price of residential natural gas and electricity prices. Each observation is a state-month. As shown in Table 1, columns 1 and 2, home energy prices are strongly positively correlated with the heating price proxy. In column 1, we include only state and month-year fixed effects. In column 2, we add our other control variables. The coefficient on the heating price proxy is less than 1 because the outcome is the average *energy* price, while the regressor is a proxy for the average *heating* price. Heating comprises roughly 40% of annual home energy use, so we would expect a 10% change in the heating price to lead to a 4% change in the home energy price, or a coefficient of 0.4. The estimated coefficient of 0.36 is quite close to this.

We next quantify how heating prices affect households' energy use and energy bills. (In principle, once we know one of these numbers, we could calculate the other, but showing both is useful given that the data are available at different geographic levels and based on different samples.) First, we examine the impact on energy usage, shown in Table 1, columns 3 and 4. As expected, higher prices lead to less energy consumption.[8] Both the outcome

[8]Appendix Table A4 shows that this cutback in usage occurs three months after the increase in the heating price, as stated in Section 3.2.

and key regressor are in logs, so the coefficient represents an elasticity. The coefficient of -0.093 implies that households cut back usage quite a bit, but not one-for-one with price. To quantify the energy-use elasticity, one needs to scale the coefficient by the corresponding price-change coefficient from columns 1 and 2.[9] We report the implied elasticity, which is -0.26, at the bottom of the table. This elasticity is similar to the winter natural gas demand elasticity for California estimated by Auffhammer and Rubin (2018) and Hahn and Metcalfe (2021). In Appendix Table A5, we show that the estimates based on our winter/non-winter specification are similar.

The elasticity having a magnitude less than 1 implies that households are spending more money on energy expenses when the heating price increases. We verify this using Census/ACS data. Columns 5 and 6 of Table 1 show that the heating price shock is associated with a 25 log point increase in energy expenses. If the result is driven by changes in winter expenses, then the coefficient is an underestimate of the impact during winter months. (We cannot isolate spending in winter because the ACS does not release the survey month, and the Census asks about annual spending on energy bills.) Columns 7 and 8 examine the outcome in levels: a 10% increase in the price of heating is associated with a $5 (in 2016 USD) increase in the monthly home energy bill, averaged over the year. To help interpret these magnitudes, note that the relative price of natural gas fell by 42% (54 log points) between 2005 and 2010. This price decline led to a 13% or $330 annual decrease in energy bills for natural gas users, using the estimates in columns 6 and 8, respectively.

To summarize, we find that households meaningfully reduce their heating use in response to an increase in their heating price, and they also experience an increase in their energy bills.[10]

[9] The relevant scale factor to convert our mortality results into an elasticity of mortality with respect to the heating price is 1; $\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt})$ incorporates information on heating sources and hence is a better proxy of the heating price than the average energy price.

[10] We also investigated the impact of heating prices on households' other non-energy expenditure patterns using the Consumer Expenditure Survey (CEX) data (Appendix Table A6). We find statistically insignificant effects, with large confidence intervals, for all broad categories of expenditure including food and alcoholic beverages; non-durable goods; and all non-energy expenditures. The effect on health expenditures is significant at 10% level.

4.2 Effect of heating price on mortality

We now turn to estimating the effect of heating prices on mortality. Table 2 shows that a higher (log) heating price increases the (log) mortality rate.[11] Column 1 reports results for all-cause mortality, controlling for all fixed effects and control variables listed earlier. The elasticity of all-cause mortality with respect to the heating price is 0.032 ($p < 0.05$).[12]

Column 2 presents results for EWM mortality. An increase in the heating price increases EWM mortality, with an elasticity of EWM mortality with respect to price of 0.059 ($p < 0.01$).[13] Given that EWM causes account for 63% of total mortality in winter, the implied elasticity of total mortality is 0.037, similar to the elasticity using all-cause mortality.

We next examine non-EWM mortality. As shown in column 3, the coefficient for the heating price proxy is very close to 0 and statistically insignificant. Non-EWM causes are, by and large, not exacerbated by exposure to cold, so the heating use channel is not applicable. However, this is not a placebo test because the non-heating consumption channel (less income to spend on non-heating expenditures) should affect non-EWM mortality. Under the assumption that the non-heating channel has similar effects on EWM and non-EWM mortality, the lack of an effect of heating prices on non-EWM mortality indicates the importance of the heating channel — changes in heating use seem to drive the effect of heating prices on mortality.

Columns 4 to 7 disaggregate the effects by broad EWM category: The overall effect on EWM mortality is mainly driven by circulatory and respiratory causes. Appendix Table A11 reports results separately for each of the 14 EWM causes. The largest effect sizes are for emphysema, other chronic lower respiratory diseases, acute myocardial infarction, and pneumonia. Interestingly, the price of heating does not exacerbate influenza mortality.

The effects we estimate are not due to deaths being moved earlier by just a short

[11] Appendix Table A7 shows the first stage of the instrumental variables regression. Appendix Tables A8, A9, and A10 show robustness to using the age-adjusted mortality rate in levels, weighting regressions by the population in 2000, and using only natural gas variation for identification.

[12] We also investigated the effect on morbidity using the Heath and Retirement Study and on hospitalizations using the National Inpatient Sample, but due to the smaller sample sizes, we were underpowered to detect even elasticities much larger than our estimated elasticity for mortality.

[13] Appendix Figure A3 shows a binned scatterplot of the relationship between EWM mortality and the instrument.

duration, or "harvesting." Appendix Table A12 shows that the cumulative mortality effect is stable in magnitude when we incorporate effects in subsequent months. (For simplicity, the table reports reduced-form estimates.) The cumulative effect is statistically significant at at least the 5 percent level when we add up to three subsequent months and marginally significant up to six months. There is not enough statistical power to determine at what point the cumulative effect becomes essentially zero. (Note that the coefficient for any specific lag is difficult to interpret because RelPrice is serially correlated and we have a finite number of months in the sample.)

We next bring in data for non-winter months to estimate the winter/non-winter specification. We use either Winter (Table 2, column 8) or HDD (column 9) to construct the additional comparison. Column 8 shows that the effect of heating prices on mortality is stronger in winter than the rest of the year. Reassuringly, the coefficient on the non-interacted heating price proxy is close to zero: the price of heating having no effect on mortality in non-winter months can be thought of as a placebo test.

Using HDD, we find that the price of heating increases mortality more in colder months. HDD is normalized so that a unit change is the difference between every day in the month being 65°F or above and being 32°F. As reported in column 9, a one-unit increase in HDD_{jt}, relative to the county's average winter HDD, leads to a 0.090 higher elasticity of EWM mortality with respect to the heating price.[14]

The results are similar but somewhat weaker when we do not control for average HDD and thus use average differences across places in the severity of their winters as additional identifying variation (see Appendix Table A13). This is consistent with previous findings that, due to adaptation (e.g., better insulated homes in colder places), atypical cold for an area is what especially affects mortality (Eurowinter Group 1997).

Appendix Tables A14 and A15 show robustness of our results to varying the definitions of winter, RelPrice, or ShareGas; excluding states with high shares of other heating fuel sources; excluding shale-gas-producing states; dropping the Great Recession period; controlling for LIHEAP, additional air pollutants, or a richer set of controls using a double-selection post-

[14]The coefficient on the heating price proxy is not interpretable because we control for the county's average winter HDD in parallel to HDD_{jt} (see Appendix C.1).

LASSO method; estimating the effects at the state level or using only within-Census division variation for identification; and varying the main set of control variables. Appendix C.2 discusses these robustness checks.

4.3 Heterogeneous effects on mortality

Table 3 augments the baseline specification to examine heterogeneous effects by poverty. Heating bills comprise a larger share of expenditures for the poor. For this reason, as well as the poor having lower baseline health and less access to health care, we expect heating prices to have larger effects on mortality among the poor. Columns 1 to 4 each use a different poverty proxy. In column 1, the proxy is whether the county's median income is in the bottom half of the distribution across counties. Columns 2 and 3 use the county's share of households below 150% of the federal poverty line, as either a continuous variable or an indicator for being below the sample median. Column 4 uses the decedent's education level, specifically an indicator for no high school degree. Across the board, the point estimates suggest larger effects among the poor, but the finding is only statistically significant in columns 2 and 3, which use the share of households below 150% of the poverty line.

Finally, Table 3, columns 5 and 6, show that the mortality effects do not significantly differ by sex or race. In Appendix C.3, we discuss heterogeneity by age groups.

5 Conclusion

This paper finds that lower heating prices reduce winter mortality. To put the estimated elasticity of all-cause mortality with respect to the price of heating of 0.032 in context, the price of natural gas relative to electricity fell by 42% between 2005 to 2010. Our findings imply that this price decline caused a 1.7% decrease in the winter mortality rate for households using natural gas for heating. Given that 58% of American households use natural gas for heating, the drop in natural gas prices reduced the US winter mortality rate by 1.0%, or, equivalently, the annual mortality rate by 0.4%. This represents 12,500 deaths per year. In terms of welfare, our results map to approximately $103 billion using a value of statistical life year of $369,000 in 2016 dollars (Kniesner and Viscusi 2019). This national-level benefit from averted deaths is twice as large as the local economic gains from fracking and should not be ignored when evaluating the effects of shale gas production (see Appendix C.4 for details). This estimate includes only relatively immediate effects, and the total benefit could be larger if there are also morbidity effects that affect mortality further out than six months. Our results suggest that reduced heating use (as opposed to other spending cutbacks households make when they face high heating bills) is the key channel through which expensive heating increases mortality.

Soaring energy prices in Europe caused by Russia's 2022 invasion of Ukraine have brought renewed attention to policies that can reduce home energy costs. Our findings highlight the health benefits of such policies. While price interventions can distort allocative efficiency, our estimates suggest the health gains from these policies can be large, particularly for low-income households.

References

Analitis, A., A. Biggeri, A. Hojs, B. Forsberg, E. Cadum, F. Ballester, J. Sunyer, K. Katsouyanni, L. Bisanti, M. Baccini, P. Michelozzi, P. Tiittanen, P. G. Goodman, and U. Kirchmayer (2008). "Effects of Cold Weather on Mortality: Results From 15 European Cities Within the PHEWE Project". *American Journal of Epidemiology* 168 (12), pp. 1397–1408.

Auffhammer, M. and E. Rubin (2018). *Natural Gas Price Elasticities and Optimal Cost Recovery Under Consumer Heterogeneity: Evidence from 300 Million Natural Gas Bills*. Working Paper 24295. National Bureau of Economic Research.

Barreca, A., K. Clay, O. Deschênes, M. Greenstone, and J. S. Shapiro (2016). "Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century". *Journal of Political Economy* 124 (1), pp. 105–159.

Barreca, A. I. and J. P. Shimshack (2012). "Absolute Humidity, Temperature, and Influenza Mortality: 30 Years of County-level Evidence from the United States". *American Journal of Epidemiology* 176 (7), S114–S122.

Bartik, A. W., J. Currie, M. Greenstone, and C. R. Knittel (2019). "The Local Economic and Welfare Consequences of Hydraulic Fracturing". *American Economic Journal: Applied Economics* 11 (4), pp. 105–55.

Beatty, T. K., L. Blow, and T. F. Crossley (2014). "Is There a 'Heat-or-Eat' Trade-off in the UK?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 177 (1), pp. 281–294.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). "Inference on treatment effects after selection among high-dimensional controls". *The Review of Economic Studies* 81 (2), pp. 608–650.

Bhattacharya, J., T. DeLeire, S. Haider, and J. Currie (2003). "Heat or Eat? Cold-Weather Shocks and Nutrition in Poor American Families". *American Journal of Public Health* 93 (7), pp. 1149–1154.

Brown, S. P. and M. K. Yücel (2008). "What Drives Natural Gas Prices?" *The Energy Journal*, pp. 45–60.

Casey, J. A., D. A. Savitz, S. G. Rasmussen, E. L. Ogburn, J. Pollak, D. G. Mercer, and B. S. Schwartz (2016). "Unconventional Natural Gas Development and Birth Outcomes in Pennsylvania, USA". *Epidemiology* 27 (2), p. 163.

Castellani, J. W. and A. J. Young (2016). "Human Physiological Responses to Cold Exposure: Acute Responses and Acclimatization to Prolonged Exposure". *Autonomic Neuroscience* 196, pp. 63–74.

Crawford, V., M McCann, and R. Stout (2003). "Changes in Seasonal Deaths from Myocardial Infarction". *QJM* 96 (1), pp. 45–52.

Critchley, R., J. Gilbertson, M. Grimsley, G. Green, and Warm Front Study Group (2007). "Living in Cold Homes after Heating Improvements: Evidence from Warm-Front, England's Home Energy Efficiency Scheme". *Applied Energy* 84 (2), pp. 147–158.

Crossley, T. F. and F. Zilio (2018). "The Health Benefits of a Targeted Cash Transfer: The UK Winter Fuel Payment". *Health Economics* 27 (9), pp. 1354–1365.

Cullen, J. A. and E. T. Mansur (2017). "Inferring Carbon Abatement Costs in Electricity Markets: A Revealed Preference Approach Using the Shale Revolution". *American Economic Journal: Economic Policy* 9 (3), pp. 106–33.

Cullen, J. B., L. Friedberg, and C. Wolfram (2004). "Consumption and Changes in Home Energy Costs: How Prevalent is the 'Heat or Eat' Decision?" University of California, Berkeley manuscript.

Currie, J., M. Greenstone, and K. Meckel (2017). "Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania". *Science Advances* 3 (12), e1603021.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008). "Physiographically Sensitive Mapping of Climatological Temperature and Precipitation Across the Conterminous United States". *International Journal of Climatology* 28 (15), pp. 2031–2064.

Davis, L. W. and L. Kilian (2011). "The allocative cost of price ceilings in the US residential market for natural gas". *Journal of Political Economy* 119 (2), pp. 212–241.

Deschênes, O. and M. Greenstone (2011). "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US". *American Economic Journal: Applied Economics* 3 (4), pp. 152–85.

Deschênes, O. and E. Moretti (2009). "Extreme Weather Events, Mortality, and Migration". *Review of Economics and Statistics* 91 (4), pp. 659–681.

Donaldson, G. C. and J. A. Wedzicha (2014). "The Causes and Consequences of Seasonal Variation in COPD Exacerbations". *International Journal of Chronic Obstructive Pulmonary Disease* 9, pp. 1101–1110.

Eccles, R. (2002). "An Explanation for the Seasonality of Acute Upper Respiratory Tract Viral Infections". *Acta Oto-Laryngologica* 122 (2), pp. 183–191.

El Ansari, W. and S. El-Silimy (2008). "Are Fuel Poverty Reduction Schemes Associated with Decreased Excess Winter Mortality in Elders? A Case Study from London, UK". *Chronic Illness* 4 (4), pp. 289–294.

Eurostat (2021). *Inability to keep home adequately warm - EU-SILC survey*. ILC_MDES01. Retrieved December 6, 2021, from https://ec.europa.eu/eurostat/databrowser/view/ILC_MDES01__custom_1485289/default/table?lang=en.

Eurowinter Group (1997). "Cold Exposure and Winter Mortality from Ischaemic Heart Disease, Cerebrovascular Disease, Respiratory Disease, and All Causes in Warm and Cold Regions of Europe". *The Lancet* 349 (9062), pp. 1341–1346.

Fell, H. and D. T. Kaffine (2018). "The Fall of Coal: Joint Impacts of Fuel Prices and Renewables on Generation and Emissions". *American Economic Journal: Economic Policy* 10 (2), pp. 90–116.

Frank, D. A., N. B. Neault, A. Skalicky, J. T. Cook, J. D. Wilson, S. Levenson, A. F. Meyers, T. Heeren, D. B. Cutts, P. H. Casey, M. M. Black, and C. Berkowitz (2006). "Heat or Eat: The Low Income Home Energy Assistance Program and Nutritional and Health Risks Among Children Less Than 3 Years of Age". *Pediatrics* 118 (5), e1293–e1302.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). "Bartik instruments: What, when, why, and how". *American Economic Review* 110 (8), pp. 2586–2624.

Green, G. and J. Gilbertson (2008). "Warm Front, Better Health: Health Impact Evaluation of the Warm Front Scheme". *Sheffield: Sheffield Hallam University*.

Grey, C. N. B., T. Schmieder-Gaite, S. Jiang, C. Nascimento, and W. Poortinga (2017). "Cold Homes, Fuel Poverty and Energy Efficiency Improvements: A Longitudinal Focus Group Approach". *Indoor and Built Environment* 26 (7), pp. 902–913.

Groundwater Protection Council (2009). *Modern Shale Gas Development in the United States: A Primer*. Research Report. Prepared for the U.S. Department of Energy, National Energy Technology Laboratory with ALL Consulting, Oklahoma City, OK.

Hahn, R. W. and R. D. Metcalfe (2021). "Efficiency and Equity Impacts of Energy Subsidies". *American Economic Review* 111 (5), pp. 1658–88.

Hartley, P. R., K. B. Medlock, and J. E. Rosthal (2008). "The Relationship of Natural Gas to Oil Prices". *The Energy Journal*, pp. 47–65.

Hausman, C. and R. Kellogg (2015). "Welfare and Distributional Implications of Shale Gas". *Brookings Papers on Economic Activity*.

Hill, E. L. (2018). "Shale Gas Development and Infant Health: Evidence from Pennsylvania". *Journal of Health Economics* 61, pp. 134–150.

Holladay, J. S. and J. LaRiviere (2017). "The Impact of Cheap Natural Gas on Marginal Emissions from Electricity Generation and Implications for Energy Policy". *Journal of Environmental Economics and Management* 85, pp. 205–227.

Howden-Chapman, P., A. Matheson, J. Crane, H. Viggers, M. Cunningham, T. Blakely, C. Cunningham, A. Woodward, K. Saville-Smith, D. O'Dea, M. Kennedy, M. Baker,

N. Waipara, R. Chapman, and G. Davie (2007). “Effect of Insulating Existing Houses on Health Inequality: Cluster Randomised Study in the Community”. *BMJ* 334 (7591), p. 460.

Hoynes, H., D. L. Miller, and J. Schaller (2012). “Who Suffers During Recessions?” *Journal of Economic Perspectives* 26 (3), pp. 27–48.

Jackson, R. B., A. Vengosh, J. W. Carey, R. J. Davies, T. H. Darrah, F. O’Sullivan, and G. Pétron (2014). “The Environmental Costs and Benefits of Fracking”. *Annual Review of Environment and Resources* 39, pp. 327–362.

Keatinge, W. R., S. R. Coleshaw, F Cotter, M Mattock, M Murphy, and R Chelliah (1984). “Increases in Platelet and Red Cell Counts, Blood Viscosity, and Arterial Pressure During Mild Surface Cooling: Factors in Mortality from Coronary and Cerebral Thrombosis in Winter”. *BMJ* 289 (6456), pp. 1405–1408.

Kniesner, T. J. and W. K. Viscusi (July 2019). *The Value of a Statistical Life*.

Knittel, C. R., K. Metaxoglou, and A. Trindade (2015). *Natural Gas Prices and Coal Displacement: Evidence from Electricity Markets*. Working Paper 21627. National Bureau of Economic Research.

Liddell, C. and C. Morris (2010). “Fuel Poverty and Human Health: a Review of Recent Evidence”. *Energy Policy* 38 (6), pp. 2987–2997.

LIHEAP Clearinghouse (2010). *Targeting LIHEAP Benefits: State Strategies Based On Household Income, Energy Burden And Heating Costs*. Tech. rep. U.S. Department of Health and Human Services.

Linn, J. and L. Muehlenbachs (2018). “The Heterogeneous Impacts of Low Natural Gas Prices on Consumers and the Environment”. *Journal of Environmental Economics and Management* 89, pp. 1–28.

Manson, S., J. Schroeder, D. Van Riper, and S. Ruggles (2017). *IPUMS National Historical Geographic Information System: Version 12.0 [Database]*. Minneapolis: University of Minnesota.

Mourtzoukou, E. and M. E. Falagas (2007). “Exposure to Cold and Respiratory Tract Infections”. *The International Journal of Tuberculosis and Lung Disease* 11 (9), pp. 938–943.

Muehlenbachs, L., E. Spiller, and C. Timmins (2015). “The Housing Market Impacts of Shale Gas Development”. *American Economic Review* 105 (12), pp. 3633–59.

Myers, E. (2019). “Are Home Buyers Inattentive? Evidence from Capitalization of Energy Costs”. *American Economic Journal: Economic Policy* 11 (2), pp. 165–88.

Neidell, M., S. Uchida, and M. Veronesi (2021). "The unintended effects from halting nuclear power production: Evidence from Fukushima Daiichi accident". *Journal of Health Economics* 79, p. 102507.

PRISM Climate Group (2004). *PRISM Gridded Climate Data*. Oregon State University.

RECS (2009). *Residential Energy Consumption Survey*. Research Report. Energy Information Agency.

— (2014). *Everywhere But Northeast, Fewer Homes Choose Natural Gas as Heating Fuel*. Research Report. Energy Information Agency.

Rostron, B. L., J. L. Boies, and E. Arias (2010). "Education Reporting and Classification on Death Certificates in the United States (Vital Health Statistics Series 2, No. 151)". *Hyattsville, MD: National Center for Health Statistics*.

Ruggles, S., S. Flood, R. Goeken, M. Schouweiler, and M. Sobek (2022). *IPUMS USA: Version 12.0 [dataset]*. Minneapolis: University of Minnesota.

Schwandt, H., J. Currie, M. Bär, J. Banks, P. Bertoli, A. Bütikofer, S. Cattan, B. Z.-Y. Chao, C. Costa, L. González, et al. (2021). "Inequality in Mortality Between Black and White Americans by Age, Place, and Cause and in Comparison to Europe, 1990 to 2018". *Proceedings of the National Academy of Sciences* 118 (40).

Snyder, R. L. and J. P. d. Melo-Abreu (2005). *Frost Protection: Fundamentals, Practice and Economics*. Vol. 1. Environment and Natural Resources Series 10. Retrieved May 2020, from http://www.fao.org/3/y7223e/y7223e00.htm. Rome: Food and Agriculture Organization of the United Nations.

Stevens, A. H., D. L. Miller, M. E. Page, and M. Filipski (2015). "The Best of Times, the Worst of Times: Understanding Pro-Cyclical Mortality". *American Economic Journal: Economic Policy* 7 (4), pp. 279–311.

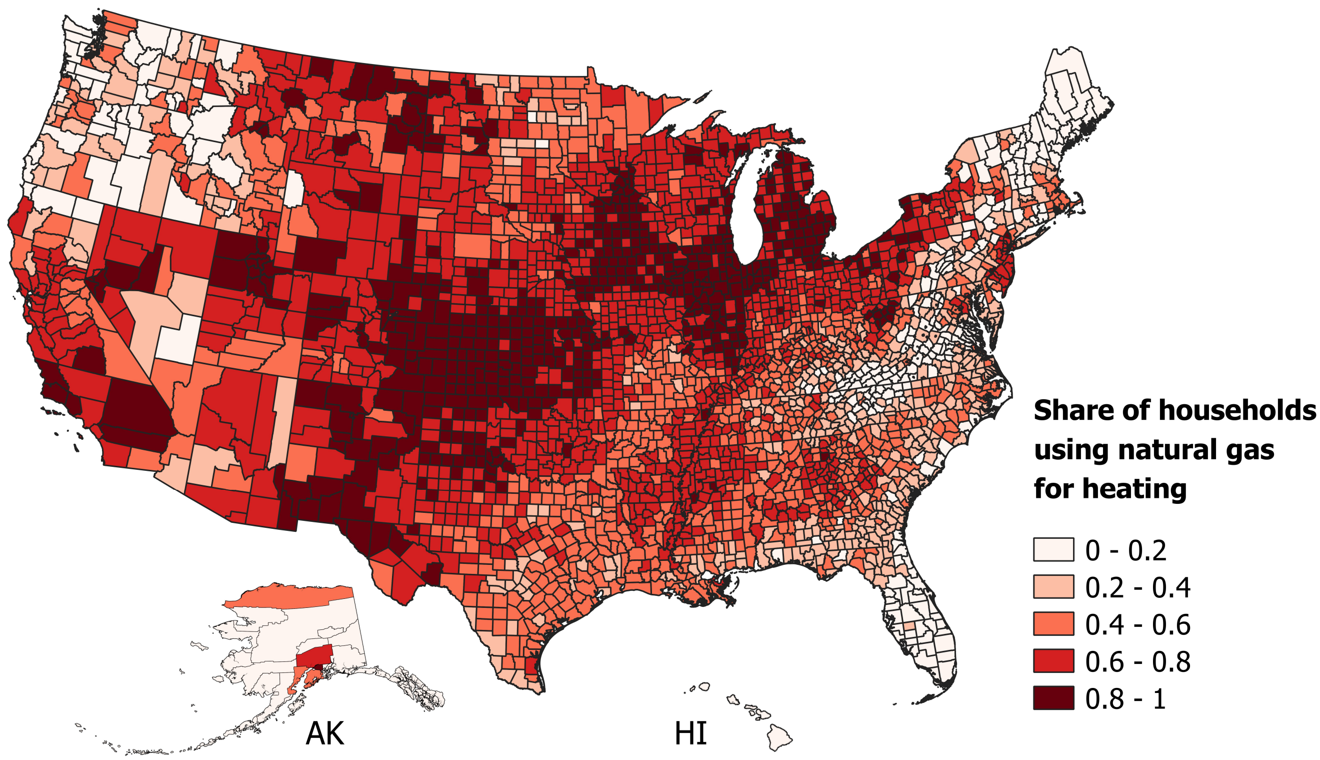
Wilkinson, P., S. Pattenden, B. Armstrong, A. Fletcher, R. S. Kovats, P. Mangtani, and A. J. McMichael (2004). "Vulnerability to Winter Mortality in Elderly People in Britain: Population Based Study". *BMJ* 329 (7467), p. 647.

Wilkinson, P., K. R. Smith, S. Beevers, C. Tonne, and T. Oreszczyn (2007). "Energy, Energy Efficiency, and the Built Environment". *The Lancet* 370 (9593), pp. 1175–1187.

Woodhouse, P., K. Khaw, M. Plummer, T. Meade, and A. Foley (1994). "Seasonal Variations of Plasma Fibrinogen and Factor VII Activity in the Elderly: Winter Infections and Death from Cardiovascular Disease". *The Lancet* 343 (8895), pp. 435–439.

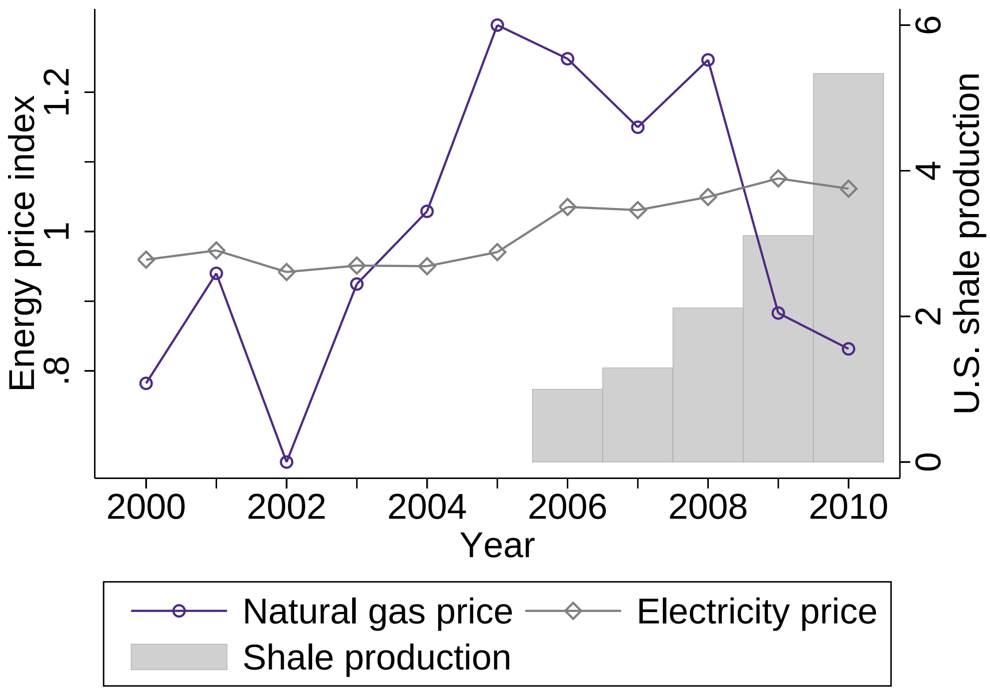
Ye, X., R. Wolff, W. Yu, P. Vaneckova, X. Pan, and S. Tong (2012). "Ambient Temperature and Morbidity: A Review of Epidemiological Evidence". *Environmental Health Perspectives* 120 (1), p. 19.

Figure 1: Share of households using natural gas for heating, by US county



Notes: The figure shows the proportion of occupied housing units in each county that report using natural gas as their main heating source. Data are from the 2000 US Census.

Figure 2: US natural gas and electricity prices, 2000 to 2010



Notes: The data series depicted with lines are the national prices of natural gas and electricity, normalized by their respective averages between 2000 and 2010 (left axis). National shale gas production in trillion cubic feet is shown as the bar chart (right axis). Data are from the US Energy Information Administration.

Table 1: Effect of heating price on energy use and energy spending

	Dependent variable:							
	Log of average electricity and gas price		Log of total energy consumption		Log of total monthly energy bill		Total monthly energy bill	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Heating price proxy	0.351***	0.361***	-0.125***	-0.0932**	0.270***	0.246***	57.4***	50.9***
	[0.0671]	[0.0700]	[0.0391]	[0.0393]	[0.0369]	[0.0352]	[7.33]	[6.94]
Observations	2,695	2,695	2,695	2,695	21,665	21,665	21,665	21,665
Mean price/quantity	21.1	21.1	22.1	22.1	220.7	220.7	220.7	220.7
Basic fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All other controls	No	Yes	No	Yes	No	Yes	No	Yes
Implied elasticity			-0.36	-0.26				

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Columns 1 to 4: The sample comprises state-year-months in the contiguous US for winter months (November–March) between 2000 and 2010. Outcomes are constructed from EIA data. Columns 5 to 8: The sample comprises county-years in the contiguous US, aggregated and crosswalked from microdata in the 2000 Census and the ACS PUMS data between 2005 and 2010. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the state-year (columns 1 to 4) or county-year (columns 5 to 8) proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the citygate price of natural gas to the residential price of electricity. Prices are state-month prices averaged over the three- and four-month lag in columns 1 to 4, and state-year prices in columns 5 to 8. *Heating price proxy* is instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$, i.e. the interaction of $ShareGas_{jt}$ in 2000 with the US-level $Log(RelPrice_{jt})$. Average electricity and gas price is the state's consumption-weighted average of the residential prices of electricity and gas, in dollars per million British Thermal Units (BTUs). Total energy consumption is the state's total delivery of natural gas and electricity to residential consumers, in trillion BTUs. Total monthly energy bill is the mean monthly bill from electricity, gas and other fuels in the county. *Basic fixed effects* are state and year-month fixed effects for columns 1 to 4, and county and year fixed effects for columns 5 to 8. *All other controls* are the interactions of $log(RelPrice_{US,t})$ with the log state or county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000, the state housing price index, the unemployment rate, the state's manufacturing sector share of total employee compensation, HDD, a quadratic in absolute humidity, the air quality indices (AQIs) for $PM_{2.5}$, PM_{10}, and NO_2, and the AQI for NO_2 squared. Implied elasticity is the ratio of the coefficient reported in that column to the corresponding coefficient from the first two columns. Monetary variables are in constant 2016 US dollars.

Table 2: Effect of heating price on mortality from all-cause and EWM causes of death

	Dependent variable: Log of mortality rate								
	All causes	All EWM causes	Non-EWM causes	Group A EWM: Non-viral, non-respiratory infections	Group G EWM: Neuro-logical diseases	Group I EWM: Circula-tory system diseases	Group J EWM: Respira-tory system diseases	All EWM causes	All EWM causes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Heating price proxy	0.032**	0.059***	0.0033	0.021	0.021	0.054**	0.099***	-0.015	0.090**
	[0.014]	[0.017]	[0.021]	[0.025]	[0.029]	[0.020]	[0.020]	[0.015]	[0.037]
Heating price proxy × Winter								0.073***	
								[0.019]	
Heating price proxy × HDD									0.090***
									[0.032]
Observations	153,296	152,927	151,113	108,659	110,742	151,589	148,583	366,668	366,668
Mean mortality rate	929.5	577.6	358.4	74.16	74.01	371.8	259.8	527.8	527.8
Months used	Winter	Winter	Winter	Winter	Winter	Winter	Winter	All	All

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US between 2000 and 2010. In columns 1 to 7, the sample is restricted to winter months (November–March). Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Winter* is a binary variable that equals one in winter months (November to March). *HDD* is the number of heating degree-days in the county for the month, based on thresholds of 65˚F, in units of ˚F-days divided by 1000, and scaled to a 30-day month. *Heating price proxy* and its interaction with *Winter/HDD* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Winter/HDD*. Columns 1 to 7: All columns control for county and year-month fixed effects, the interactions of $log(RelPrice_{US,t})$ with the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000, the state housing price index, the unemployment rate, the state's manufacturing sector share of total employee compensation, HDD, a quadratic in absolute humidity, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, and the AQI for NO_2 squared. Columns 8 and 9: All columns control for the above set plus the following: all possible two-way interactions between $ShareGas_{j,2000}$, $log(RelPrice_{US,t})$, and *Winter/HDD*; and the two- and three-way interactions among $log(RelPrice_{US,t})$, *Winter/HDD*, and each of the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000. Column 9 also includes the interaction of the average county HDD in winter months with $log(RelPrice_{US,t})$; and the three-way interactions of the average county HDD in winter months, $log(RelPrice_{US,t})$, and each of $ShareGas_{j,2000}$, the log county household income in 1999, and the share of people aged 70 and above in 2000.

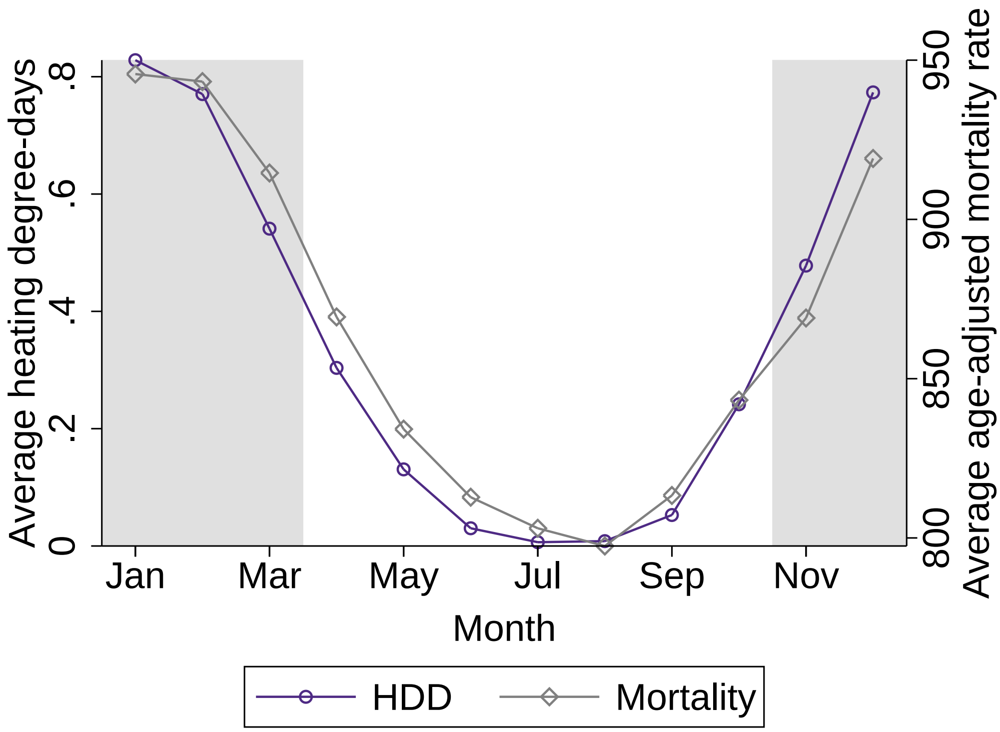
Table 3: Heterogeneous effects on mortality

	Dependent variable: Log of all-EWM-causes mortality rate Trait is:					
	Below-median county income	Proportion below 150% of poverty line	Above-median proportion below 150% of poverty line	No high school degree	Male	Black
	(1)	(2)	(3)	(4)	(5)	(6)
Heating price proxy × Trait	0.021	0.36**	0.057**	0.033	0.013	0.013
	[0.032]	[0.17]	[0.026]	[0.039]	[0.026]	[0.044]
Heating price proxy	0.049***	-0.025	0.038**	0.027	0.058***	0.053***
	[0.016]	[0.037]	[0.016]	[0.045]	[0.017]	[0.017]
Observations	152,927	152,927	152,927	284,700	300,311	218,275
Mean mortality rate	577.6	577.6	577.6	999.4	605.3	739.4
Implied effect for Trait = 1	0.07**	0.33**	0.10***	0.06	0.07***	0.07
	[0.03]	[0.14]	[0.03]	[0.05]	[0.02]	[0.04]

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. For columns 1 to 3, the sample comprises county-year-months in the contiguous US for winter months (November–March) between 2000 and 2010. For columns 4, 5, and 6, the sample comprises county-year-months-education, county-year-months-sex, and county-year-months-race groups, respectively, for winter months. Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the log of the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. Column 1: *Trait* is an indicator variable that equals one if the county's median household income is below the median of all counties in the sample in 1999. Column 2: *Trait* is the proportion of households in the county with income in 1999 below 150 percent of the poverty threshold. Column 3: *Trait* is an indicator variable that equals one if the proportion from column 2 is above the median of all counties in the sample. Column 4: *Trait* is an indicator variable that equals one for the subgroup that did not complete high school. Column 5: *Trait* is an indicator variable that equals one for the male population. Column 6: *Trait* is an indicator variable that equals one for the Black population; non-Black and non-White populations are excluded from the sample. *Heating price proxy* and its interaction with *Trait* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Trait*. All columns include all fixed effects and control variables from column 2 of Table 2, the main effect for *Trait*, and the interaction of each fixed effect or control variable with *Trait*.

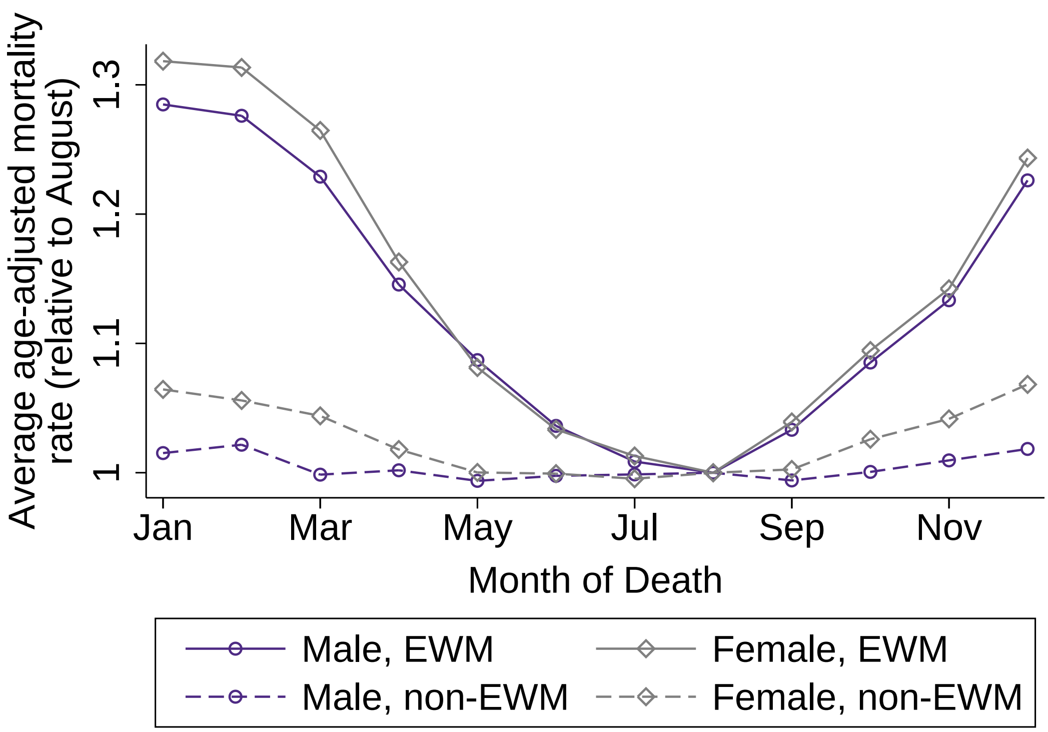
A Appendix figures and tables

Appendix Figure A1: Heating degree-days and monthly mortality in the US



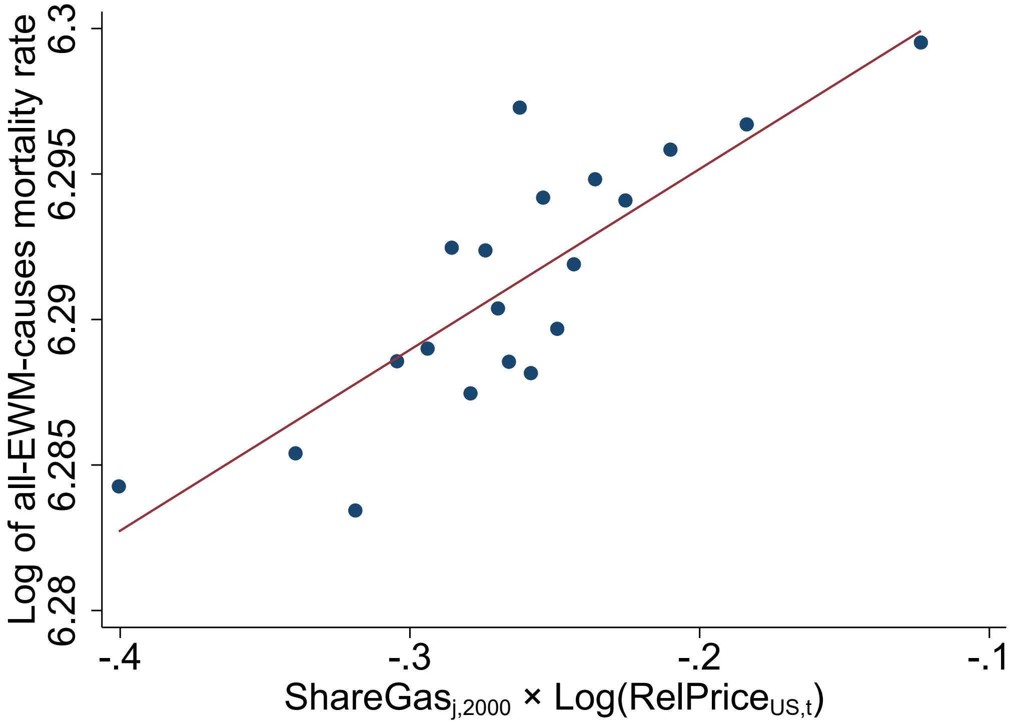
Notes: Average heating degree-days (HDD) and average age-adjusted mortality rates across US counties (excluding Hawaii and Alaska) between 2000 and 2010 plotted by month. Average HDD is computed using temperature data from the PRISM Climate Group, and is based on a threshold of 65°F, in units of °F-days divided by 1000, and scaled to a 30-day month. Average age-adjusted mortality rates are computed using the NCHS mortality data and expressed per 100,000 population on an annualized basis. Months we define as winter in our analysis (November–March) are shaded in the background.

Appendix Figure A2: Seasonality in mortality for EWM and non-EWM causes



Notes: Average age-adjusted mortality rates across US counties (excluding Hawaii and Alaska) between 2000 and 2010, broken down by sex and EWM versus other causes. EWM causes are those that exhibit a strong pattern of higher mortality in winter than the rest of the year, as described in the text; see data appendix for further details. We normalize each series by its value in August (the month with the lowest all-cause mortality rate). Age-adjusted mortality rates are computed using the NCHS mortality data.

Appendix Figure A3: Binned scatterplot of the relationship between EWM mortality and the heating price



Notes: The sample comprises county-year-months in the contiguous US for winter months (November–March) between 2000 and 2010. Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ is the instrument used in Table 2, where $ShareGas_{j,2000}$ is the proportion of occupied housing units in the county in 2000 with natural gas as their main heating source, and $RelPrice_{US,t}$ is the ratio of the monthly citygate price of natural gas in the US, averaged over the three- and four-month lag, to the corresponding residential price of electricity. All fixed effects and control variables from Table 2, column 2 are partialled out before plotting.

Appendix Table A1: Regression of $ShareGas_{jt}$ on $\log(RelPrice_{jt})$

	Dependent variable: ShareGas					
	0 lags (contem-poraneous)	1 lag	2 lags	3 lags	4 lags	5 lags
	(1)	(2)	(3)	(4)	(5)	(6)
Log(RelPrice)	0.016	0.0023	-0.0097	-0.0071	-0.010	-0.016
	[0.013]	[0.014]	[0.013]	[0.013]	[0.012]	[0.012]
Observations	539	539	539	539	539	539

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises state-years in the contiguous US between 2000 and 2010. The dependent variable, *ShareGas* is the proportion of occupied housing units in the state with natural gas as their main heating source, aggregated from Census and ACS microdata to the state-year level. The main regressor, *Log(RelPrice)*, is the log of the ratio of the state-year citygate price of natural gas to residential price of electricity at the specified lag. All columns include state and year fixed effects.

Appendix Table A2: Relationship between state-level *RelPrice* and unemployment rate

	Dependent variable:		
	$\text{Log}(\text{RelPrice}_{\text{jt}})$	$\text{Log}(\text{RelPrice}_{\text{jt}})$	Δ $\text{Log}(\text{RelPrice}_{\text{jt}})$, 2005–2010
	(1)	(2)	(3)
Unemployment rate	-0.014	-0.0040	
	[0.011]	[0.0091]	
Δ Unemployment rate, 2005–2010			-0.0011
			[0.016]
Sample years	2000–2010	2005–2010	Cross-sectional
Observations	6336	3456	48
Implied 2005–2010 log point change in RelPrice	-5.6	-1.5	-0.4
Percentage of observed RelPrice decrease (%)	10.4	2.9	0.8

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises state-year-months in the contiguous US between 2000 and 2010 in column 1, state-year-months between 2005 and 2010 in column 2, and states in column 3. *Log(RelPrice)* is the log of the ratio of the state's monthly citygate price of natural gas to the state's monthly residential price of electricity in columns 1 and 2; it is similarly defined in column 3 except based on annual prices. Δ *Log(RelPrice)* or Δ unemployment rate is the change in the variable between the two years specified, i.e. 2005 to 2010. The unemployment rate is the state's monthly (columns 1 and 2) or annual (column 3) unemployment rate. Implied 2005 to 2010 log point change in *RelPrice* is 100 times the coefficient times the change from 2005 to 2010 in the average unemployment rate. The average unemployment rates among all states and months in 2005 and 2010 are 4.9% and 8.8% respectively. In the next row, we report this percentage divided by the actual decrease in national *RelPrice* of 54 log points. Columns 1 and 2 include state and year-month fixed effects. No additional control variables are included in column 3.

Appendix Table A3: Causes of death exhibiting high excess winter mortality

Cause of death (ICD-10 codes)	Mean monthly mortality rate	Level coefficient	Log coefficient
Septicemia (A40-A41)	0.95	0.14	0.14
Parkinson's disease (G20-G21)	0.53	0.08	0.16
Alzheimer's disease (G30)	1.92	0.36	0.18
Acute myocardial infarction (I21-I22)	4.34	0.62	0.14
All other forms of chronic ischemic heart disease (I20, I25.1-I25.9)	6.32	0.80	0.12
Heart failure (I50)	1.61	0.21	0.13
Cerebrovascular diseases (I60-I69)	4.12	0.52	0.12
Atherosclerosis (I70)	0.30	0.04	0.14
Influenza (J09-J11)	0.04	0.06	2.21
Pneumonia (J12-J18)	1.63	0.58	0.34
Emphysema (J43)	0.38	0.08	0.21
Other chronic lower respiratory diseases (J44, J47)	3.11	0.63	0.20
Pneumonitis due to solids and liquids (J69)	0.47	0.09	0.18
Other diseases of respiratory system (J00-J06, J30-J39, J67, J70-J98)	0.77	0.11	0.14
All other diseases (Residual)*	6.17	0.80	0.13
Accidental exposure to smoke, fire and flames (X00-X09)*	0.09	0.05	0.56

Notes: Mortality rates are expressed per 100,000 population and computed using the NCHS mortality data. The 75th percentile of level and log coefficient are 0.02 and 0.12, respectively. We remove *All other diseases* and *Accidental exposure to smoke, fire and flames* (marked with *) when we analyze mortality from high-EWM causes. See the data appendix for further details on the selection of high-EWM causes of deaths.

Appendix Table A4: Effect of heating price on energy consumption at various lags of *RelPrice*

	Dependent variable: Log of total energy consumption						
	0 lags (contemporaneous)	1 lag	2 lags	3 lags	4 lags	5 lags	6 lags
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Heating price proxy	-0.047	-0.024	-0.018	-0.069*	-0.11**	-0.12***	-0.12***
	[0.038]	[0.037]	[0.037]	[0.037]	[0.043]	[0.040]	[0.039]
Observations	2,695	2,695	2,695	2,695	2,695	2,694	2,694
Mean quantity	22.1	22.1	22.1	22.1	22.1	22.1	22.1

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises state-year-months in the contiguous US for winter months (November–March) between 2000 and 2010. Total energy consumption is the state's total delivery of natural gas and electricity to residential consumers, in trillion BTUs. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the state-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the monthly citygate price of natural gas in the state-month, lagged by the number of months indicated in each column, to the corresponding residential price of electricity. *Heating price proxy* is instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$, where $RelPrice_{US,t}$ is similarly lagged by the number of months indicated in each column. Monetary variables are in constant 2016 US dollars. All columns include all fixed effects and control variables from column 4 of Table 1.

Appendix Table A5: Winter/non-winter specification estimates of effects on average energy price and consumption

	Dependent variable: Log of average electricity and gas price		Dependent variable: Log of total energy consumption	
	(1)	(2)	(3)	(4)
Heating price proxy × Winter	0.32***		-0.098**	
	[0.045]		[0.048]	
Heating price proxy × HDD		0.31***		-0.059
		[0.070]		[0.076]
Observations	6,468	6,468	6,468	6,468
Mean price/quantity	25.5	25.5	16.0	16.0
Implied elasticity			-0.31	-0.19

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises state-year-months in the contiguous US between 2000 and 2010. Average electricity and gas price is the state's consumption-weighted average of the residential prices of electricity and gas, in dollars per million BTUs. Total energy consumption is the state's total delivery of natural gas and electricity to residential consumers, in trillion BTUs. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the state-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Winter* is a binary variable that equals one in winter months (November to March). *HDD* is the number of heating degree-days in the county for the month, based on a threshold of 65°F, in units of °F-days divided by 1000, and scaled to a 30-day month. *Heating price proxy* and its interaction with *Winter/HDD* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Winter/HDD*. Monetary variables are in constant 2016 US dollars. Implied elasticity is the ratio of the coefficient reported in that column to the corresponding coefficient from the first two columns. All columns include fixed effects and control variables analogous to those used in columns 8 and 9 of Table 2.

Appendix Table A6: Effect of heating price on non-energy expenditures in the CEX

	Dependent variable: IHS of expenditure on:				
	Food & alcohol (1)	Strictly non-durable (2)	Non-durable (3)	All non-utilities (4)	Health (5)
Heating price proxy	-0.046	0.0099	0.022	0.011	0.36*
	[0.062]	[0.051]	[0.047]	[0.048]	[0.20]
Observations	53,938	53,894	53,921	53,935	53,691
Mean expenditure	595.2	973.2	1,364.8	2,751.1	203.5

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises households on the contiguous US interviewed between February and April and between 2000 and 2010 in the CEX Interview Survey. Expenditures are average monthly expenditures for the three months prior to interview, in constant 2016 dollars. IHS refers to the inverse hyperbolic sine transformation. "Strictly non-durable" expenditures includes food and alcohol, and CEX categories like household operations, gas, personal care, and tobacco. "Non-durable expenditures" additionally includes semi-durable categories like apparel, health and reading materials. "All non-utilities expenditures" also includes durable expenditures such as home furnishings, entertainment equipment, and auto purchases. Utilities expenditures are excluded from all categories. *Heating price proxy* is the state-level analog of that in Table 2, with an additional averaging over the three months prior to the interview to match the interview structure of the data. The specification uses the instrument, fixed effects, and control variables from column 2 of Table 2, with the same additional averaging to account for the interview structure of the data. The specification additionally includes the following household-level controls: family size, indicators for any household member aged over 64 or under 18, whether the reference person does not have a high school degree, the age of the reference person, indicators for race categories (Black, non-Black Hispanic, and others), and log household income.

Appendix Table A7: First stage of instrumental variables regression

	Dependent variable: $\text{ShareGas}_{\text{jt}} \times \text{Log}(\text{RelPrice}_{\text{jt}})$ (heating price proxy)	
	(1)	(2)
$\text{ShareGas}_{\text{j,2000}} \times \text{Log}(\text{RelPrice}_{\text{US,t}})$	1.07***	1.06***
	[0.059]	[0.050]
Observations	153,340	153,340
F statistic on instrument	329.8	450.1
Basic fixed effects	No	Yes
All other controls	Yes	Yes

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US for winter months (November–March) between 2000 and 2010. $ShareGas_{jt} \times Log(RelPrice_{jt})$ is the heating price proxy used in our main specification, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ is the instrument used in our main specification, where $ShareGas_{j,2000}$ is the proportion of occupied housing units in the county in 2000 with natural gas as their main heating source, and $RelPrice_{US,t}$ is the ratio of the monthly citygate price of natural gas in the US, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Basic fixed effects* are county and year-month fixed effects. *All other controls* are the interactions of *log(RelPrice)* with the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000, the state housing price index, the unemployment rate, the state's manufacturing sector share of total employee compensation, HDD, a quadratic in absolute humidity, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, and the AQI for NO_2 squared.

Appendix Table A8: Effect of heating price on mortality using mortality rate in levels

	Dependent variable: Mortality rate								
	All causes	All EWM causes	Non-EWM causes	Group A EWM: Non-viral, non-respiratory infections	Group G EWM: Neuro-logical diseases	Group I EWM: Circula-tory system diseases	Group J EWM: Respira-tory system diseases	All EWM causes	All EWM causes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Heating price proxy	33.3***	31.8***	1.53	0.31	-1.88	17.8**	22.7***	-2.96	34.9*
	[12.4]	[9.66]	[7.24]	[2.67]	[2.37]	[7.21]	[5.17]	[6.28]	[19.4]
Heating price proxy × Winter								34.0***	
								[9.81]	
Heating price proxy × HDD									38.7**
									[15.8]
Observations	153,340	153,340	153,340	153,340	153,340	153,340	153,340	368,016	368,016
Mean mortality rate	929.2	576.0	353.2	52.55	53.45	367.5	251.7	525.9	525.9
Months used	Winter	Winter	Winter	Winter	Winter	Winter	Winter	All	All

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US between 2000 and 2010. In columns 1 to 7, the sample is restricted to winter months (November–March). Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Winter* is a binary variable that equals one in winter months (November to March). *HDD* is the number of heating degree-days in the county for the month, based on thresholds of 65˚F, in units of ˚F-days divided by 1000, and scaled to a 30-day month. *Heating price proxy* and its interaction with *Winter/HDD* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Winter/HDD*. Columns 1 to 7: All columns control for county and year-month fixed effects, the interactions of $log(RelPrice_{US,t})$ with the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000, the state housing price index, the unemployment rate, the state's manufacturing sector share of total employee compensation, HDD, a quadratic in absolute humidity, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, and the AQI for NO_2 squared. Columns 8 and 9: All columns control for the above set plus the following: all possible two-way interactions between $ShareGas_{j,2000}$, $log(RelPrice_{US,t})$, and *Winter/HDD*; and the two- and three-way interactions among $log(RelPrice_{US,t})$, *Winter/HDD*, and each of the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000. Column 9 also includes the interaction of the average county HDD in winter months with $log(RelPrice_{US,t})$; and the three-way interactions of the average county HDD in winter months, $log(RelPrice_{US,t})$, and each of $ShareGas_{j,2000}$, the log county household income in 1999, and the share of people aged 70 and above in 2000.

Appendix Table A9: Effect of heating price on mortality, weighted by population in 2000

	Dependent variable: Log of mortality rate								
	All causes	All EWM causes	Non-EWM causes	Group A EWM: Non-viral, non-respiratory infections	Group G EWM: Neuro-logical diseases	Group I EWM: Circula-tory system diseases	Group J EWM: Respira-tory system diseases	All EWM causes	All EWM causes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Heating price proxy	0.029***	0.046***	0.0091	0.0020	0.025	0.049***	0.055***	0.0054	0.036
	[0.0087]	[0.012]	[0.010]	[0.027]	[0.043]	[0.014]	[0.018]	[0.011]	[0.023]
Heating price proxy × Winter								0.037***	
								[0.011]	
Heating price proxy × HDD									0.036*
									[0.021]
Observations	153,296	152,927	151,113	108,659	110,742	151,589	148,583	366,668	366,668
Mean mortality rate	864.2	527.3	337.5	58.11	52.46	330.9	229.2	483.2	483.2
Months used	Winter	Winter	Winter	Winter	Winter	Winter	Winter	All	All

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US between 2000 and 2010. In columns 1 to 7, the sample is restricted to winter months (November–March). Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Winter* is a binary variable that equals one in winter months (November to March). *HDD* is the number of heating degree-days in the county for the month, based on thresholds of 65˚F, in units of ˚F-days divided by 1000, and scaled to a 30-day month. *Heating price proxy* and its interaction with *Winter/HDD* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Winter/HDD*. Columns 1 to 7: All columns control for county and year-month fixed effects, the interactions of $log(RelPrice_{US,t})$ with the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000, the state housing price index, the unemployment rate, the state's manufacturing sector share of total employee compensation, HDD, a quadratic in absolute humidity, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, and the AQI for NO_2 squared. Columns 8 and 9: All columns control for the above set plus the following: all possible two-way interactions between $ShareGas_{j,2000}$, $log(RelPrice_{US,t})$, and *Winter/HDD*; and the two- and three-way interactions among $log(RelPrice_{US,t})$, *Winter/HDD*, and each of the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000. Column 9 also includes the interaction of the average county HDD in winter months with $log(RelPrice_{US,t})$; and the three-way interactions of the average county HDD in winter months, $log(RelPrice_{US,t})$, and each of $ShareGas_{j,2000}$, the log county household income in 1999, and the share of people aged 70 and above in 2000. All columns are weighted by the county population in 2000.

Appendix Table A10: Effect of heating price on mortality, using only natural gas price variation

	Dependent variable: Log of mortality rate								
	All causes	All EWM causes	Non-EWM causes	Group A EWM: Non-viral, non-respiratory infections	Group G EWM: Neuro-logical diseases	Group I EWM: Circula-tory system diseases	Group J EWM: Respira-tory system diseases	All EWM causes	All EWM causes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Heating price proxy	0.023	0.050***	-0.0059	-0.00079	0.0069	0.041*	0.087***	-0.017	0.069*
	[0.015]	[0.018]	[0.022]	[0.022]	[0.029]	[0.021]	[0.021]	[0.015]	[0.040]
Heating price proxy × Winter								0.068***	
								[0.019]	
Heating price proxy × HDD									0.066**
									[0.026]
Observations	153,296	152,927	151,113	108,659	110,742	151,589	148,583	366,668	366,668
Mean mortality rate	929.5	577.6	358.4	74.16	74.01	371.8	259.8	527.8	527.8
Months used	Winter	Winter	Winter	Winter	Winter	Winter	Winter	All	All

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US between 2000 and 2010. In columns 1 to 7, the sample is restricted to winter months (November–March). Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag ($GasPrice_{jt}$), to the corresponding residential price of electricity ($ElecPrice_{jt}$). *Winter* is a binary variable that equals one in winter months (November to March). *HDD* is the number of heating degree-days in the county for the month, based on thresholds of 65˚F, in units of ˚F-days divided by 1000, and scaled to a 30-day month. *Heating price proxy* and its interaction with *Winter/HDD* are instrumented using $ShareGas_{j,2000} \times Log(GasPrice_{US,t})$ and its interaction with *Winter/HDD*. Columns 1 to 7: All columns control for county and year-month fixed effects, $ShareGas_{j,2000} \times Log(ElecPrice_{US,t})$, the interactions of $log(GasPrice_{US,t})$ with the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000, the state housing price index, the unemployment rate, the state's manufacturing sector share of total employee compensation, HDD, a quadratic in absolute humidity, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, and the AQI for NO_2 squared. Columns 8 and 9: All columns control for the above set plus the following: $ShareGas_{j,2000} \times Log(ElecPrice_{US,t})$; $ShareGas_{j,2000} \times Log(ElecPrice_{US,t}) \times Winter/HDD$; all possible two-way interactions between $ShareGas_{j,2000}$, $log(GasPrice_{US,t})$, and *Winter/HDD*; and the two- and three-way interactions among $log(GasPrice_{US,t})$, *Winter/HDD*, and each of the log county household income in 1999 (25th, 50th, and 75th percentiles) and the share of people aged 70 and above in 2000. Column 9 also includes the interaction of the average county HDD in winter months with $log(GasPrice_{US,t})$; and the three-way interactions of the average county HDD in winter months, $log(GasPrice_{US,t})$, and each of $ShareGas_{j,2000}$, the log county household income in 1999, and the share of people aged 70 and above in 2000. All columns are weighted by the county population in 2000.

Appendix Table A11: Effect of heating price on mortality, by specific cause of death

Dependent variable: Log of specified disease mortality rate			
Septicemia	0.021 [0.025] {74.2}	Atherosclerosis	0.053 [0.044] {45.9}
Parkinson's disease	0.044 [0.026] {32.2}	Influenza	-0.14 [0.12] {24.4}
Alzheimer's disease	0.030 [0.031] {63.2}	Pneumonia	0.10*** [0.031] {104.9}
Acute myocardial infarction	0.11*** [0.031] {107.3}	Emphysema	0.13*** [0.044] {29.7}
Chronic ischemic heart disease	0.080*** [0.027] {158.0}	Other chronic lower respiratory diseases	0.11*** [0.023] {114.2}
Heart failure	0.055** [0.023] {137.4}	Pneumonitis (solids and liquids)	0.053 [0.042] {44.4}
Cerebrovascular diseases	0.082** [0.031] {114.4}	Other respiratory diseases	0.053* [0.028] {107.4}

Notes: Each cell shows the result from a separate regression, and reports the coefficient on *Heating price proxy*, the corresponding standard error clustered by state in square brackets, and the mean mortality rate of the specified cause in curly brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US for winter months (November–March) between 2000 and 2010. Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Heating price proxy* is instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$. All columns include all fixed effects and control variables from column 2 of Table 2.

Appendix Table A12: Dynamic effects of heating price on mortality

	Dependent variable: Log of all-EWM-causes mortality rate						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Contemporaneous reduced form effect	0.062***	0.050	0.12**	0.047	0.049	0.034	0.048
	[0.019]	[0.051]	[0.051]	[0.046]	[0.051]	[0.055]	[0.052]
Effect on mortality 1 month after		0.0019	-0.16**	-0.027	-0.035	-0.0011	-0.00094
		[0.050]	[0.078]	[0.082]	[0.086]	[0.098]	[0.091]
Effect on mortality 2 months after			0.11*	-0.058	-0.026	-0.055	-0.093
			[0.059]	[0.089]	[0.10]	[0.11]	[0.100]
Effect on mortality 3 months after				0.13**	0.052	0.12	0.18*
				[0.057]	[0.097]	[0.11]	[0.10]
Effect on mortality 4 months after					0.014	-0.11	-0.18
					[0.049]	[0.099]	[0.11]
Effect on mortality 5 months after						0.097*	0.16
						[0.057]	[0.099]
Effect on mortality 6 months after							-0.036
							[0.053]
Observations	152,927	183,510	214,043	244,552	275,071	305,602	336,113
Cumulative effect	0.06***	0.05**	0.07***	0.09***	0.05*	0.08**	0.08*
	[0.02]	[0.02]	[0.03]	[0.03]	[0.03]	[0.04]	[0.04]

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Column 1 shows the reduced-form analog of our main estimates, and subsequent columns show dynamic reduced-form effects. The sample comprises county-year-months in the contiguous US between 2000 and 2010. The sample is restricted to months November to March in column 1, November to April in column 2, November to May in column 3, November to June in column 4, November to July in column 5, November to August in column 6, and November to September in column 7. The specification used is $log(m_{jt}) = \sum_{k=0}^{K} \beta_k ShareGas_{j,2000} \times log(RelPrice_{US,t-k}) \times MonthofEffect_k + \gamma_k ShareGas_{j,2000} \times log(RelPrice_{US,t-k}) + Controls + \epsilon_{jt}$, where $MonthofEffect_0$ takes on a value of one in the months of November to March; $MonthofEffect_1$ takes on a value of one in the months of December to April; $MonthofEffect_2$ takes on a value of one in the months of January to May; $MonthofEffect_3$ takes on a value of one in the months of February to June; $MonthofEffect_4$ takes on a value of one in the months of March to July; $MonthofEffect_5$ takes on a value of one in the months of April to August; $MonthofEffect_6$ takes on a value of one in the months of May to September; and $Controls$ are all fixed effects and control variables from column 2 of Table 2 and are fully interacted with the $MonthofEffect_k$ dummies. K, the total number of months after the contemporaneous effect, is 0 in column 1, 1 in column 2, and so on. The coefficients shown are β_k's, the effect of the winter price of heating on winter mortality k months after winter, after accounting for intertemporal correlation (since we estimate the β_k's jointly), and after removing the effect on mortality in irrelevant months through $MonthofEffect_k$ dummies (e.g., April is not a winter month, so is not relevant for the contemporaneous effect). *Cumulative effect* is the sum of the β_k's. All other definitions not noted above are as in column 2 of Table 2.

Appendix Table A13: Winter/non-winter specification mortality estimates

	Dependent variable: Log of mortality rate				
	All causes	All EWM causes	Group I EWM: Circulatory system diseases	Group J EWM: Respiratory system diseases	Non-EWM causes
	(1)	(2)	(3)	(4)	(5)
Panel A: Winter/non-winter specification using winter.					
Heating price proxy	-0.0088	-0.015	0.0082	-0.0060	0.0013
	[0.0096]	[0.015]	[0.016]	[0.020]	[0.013]
Heating price proxy × Winter	0.039**	0.073***	0.043**	0.10***	-0.00089
	[0.015]	[0.019]	[0.019]	[0.027]	[0.022]
Panel B: Winter/non-winter specification using HDD.					
Heating price proxy	0.054*	0.090**	0.095**	0.088*	0.035
	[0.028]	[0.037]	[0.036]	[0.046]	[0.035]
Heating price proxy × HDD	0.043*	0.090***	0.065*	0.10**	-0.0055
	[0.024]	[0.032]	[0.034]	[0.039]	[0.029]
Panel C: Without controlling in parallel for average winter HDD.					
Heating price proxy	-0.0038	-0.0072	0.014	0.0042	0.0045
	[0.010]	[0.018]	[0.020]	[0.021]	[0.015]
Heating price proxy × HDD	0.033*	0.058**	0.033	0.074**	0.0040
	[0.018]	[0.027]	[0.030]	[0.034]	[0.021]
Observations	367,905	366,668	362,930	353,692	362,545
Mean mortality rate	872.6	527.8	343.5	232.7	351.7

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months in the contiguous US between 2000 and 2010. Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *Winter* is a binary variable that equals one in winter months (November to March). *HDD* is the number of heating degree-days in the county for the month, based on thresholds of 65˚F, in units of ˚F-days divided by 1000, and scaled to a 30-day month. *Heating price proxy* and its interaction with *Winter/HDD* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Winter/HDD*. All columns in panels A and B include the fixed effects and control variables from columns 8 and 9 respectively of Table 2. Panel C excludes from this set of control variables the two-way and three-way interactions based on the county's average HDD in winter months.

Appendix Table A14: Effect of heating price on mortality: Robustness checks

		Dependent variable: Log of all-EWM-causes mortality rate		
		Baseline specification	Winter/non-winter specification using winter	Winter/non-winter specification using HDD
		(1)	(2)	(3)
1	Preferred specification	0.059***	0.073***	0.090***
		[0.017]	[0.019]	[0.032]
2	Winter defined as December to March	0.050**	0.065***	n/a
		[0.019]	[0.020]	
3	Winter defined as December to February	0.052**	0.074***	n/a
		[0.024]	[0.027]	
4	Using previous non-winter as comparison group	n/a	0.093***	0.090**
			[0.024]	[0.036]
5	Use residential gas price, averaged over 2nd and 3rd lags	0.054**	0.039	0.065
		[0.025]	[0.030]	[0.048]
6	Use annual residential gas price	0.084**	0.095***	0.065
		[0.033]	[0.032]	[0.052]
7	ShareGas defined as $gas/(gas + electricity)$	0.045**	0.066***	0.082**
		[0.017]	[0.020]	[0.039]
8	Exclude states with share of gas or electricity < 75%	0.058***	0.072***	0.090**
		[0.019]	[0.021]	[0.043]
9	Exclude fracking states	0.055***	0.068***	0.097***
		[0.018]	[0.018]	[0.030]
10	Exclude Great Recession	0.040**	0.066***	0.083**
		[0.019]	[0.019]	[0.032]
11	Control for Log(LIHEAP per capita)	0.058***	0.073***	0.090***
		[0.018]	[0.019]	[0.032]
12	Control for all pollutants	0.057***	0.073***	0.091***
		[0.018]	[0.019]	[0.032]
13	Controls selected by double-selection post-Lasso method	0.059***	0.063***	0.074**
		[0.018]	[0.019]	[0.031]
14	State-level regression	0.090***	0.068***	0.100**
		[0.021]	[0.020]	[0.043]
15	State-level regression, using only within-division variation	0.097**	0.047	0.11
		[0.040]	[0.041]	[0.093]
16	State-level regression, using annual price variation	0.097***	0.084***	0.095**
		[0.027]	[0.025]	[0.042]

Notes: Each cell shows the result from a separate regression, and reports the coefficient on *Heating price proxy* (column 1), *Heating price proxy* × *Winter* (column 2), or *Heating price proxy* × *HDD* (column 3). The corresponding standard error, clustered by state, is shown in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Row 1 repeats results from our preferred specifications in columns 2, 8 and 9 of Table 2 respectively. Each row from 2-12 shows a change in specification compared to row 1. Row 2: The sample excludes November, and in column 2 uses December to March as winter months. Row 3 : The sample excludes November and March, and in column 2 uses December to February as winter months. Rows 2 and 3, column 3: The winter/non-winter specification using HDD is the same as in Table 2 since HDD is defined independently of winter. Row 4: Winter/non-winter specification additionally includes $ShareGas_{j,2000} \times log(RelPrice_{US,t}) \times HeatingYearFE_t$ and $ShareGas_{j,2000} \times HeatingYearFE_t$, where $HeatingYearFE_t$ are fixed effects for the months from April of one year to March of the next. Row 5: $RelPrice_{jt}$ is constructed as the ratio of the monthly residential price of natural gas in the state, averaged over the two- and three-month lag, to the corresponding residential price of electricity. $RelPrice_{US,t}$ is similarly constructed. Row 6: $RelPrice_{jt}$ is constructed as the ratio of the annual residential price of natural gas in the state to the corresponding residential price of electricity. $RelPrice_{US,t}$ is similarly constructed. Row 7: $ShareGas_{jt}$ is the number of occupied housing units in the county with natural gas as their main heating source as a proportion of the number with natural gas and electricity. $ShareGas_{j,2000}$ is similarly modified. Row 8: The sample excludes ME, VT, NH, CT, RI, MA, NY, PA, and DE, which are the states in which the share of households using gas or electricity for heating is less than 75 %. Row 9: The sample excludes AR, LA, ND, OK, PA, TX, and WV, which are the states with significant production of shale natural gas. Row 10: The sample excludes months between December 2007 and June 2009, inclusive. This is the period of the Great Recession as defined by the NBER Business Cycle Dating Committee. Row 11: The specification includes the log of total LIHEAP assistance funds per capita in the state-fiscal year. Row 12: The specification includes the AQIs of carbon monoxide, ozone, and sulfur dioxide as control variables. Row 13: The specification includes controls selected by the double-post-LASSO method (Belloni et al. 2014) – see data appendix for further details. Row 14: Regressions are at the state-month level. Row 15: Regressions are at the state-month level. Column 1 adds Census division fixed effects interacted with *log(RelPrice)*. Columns 2 and 3 control for all possible two- and three-way interactions among Census division, *log(RelPrice)*, and *Winter*/*HDD*. Column 3 also includes all possible two- and three-way interactions among Census division, *log(RelPrice)*, and average state HDD. Row 16: Regressions are at the state-month level. $RelPrice_{jt}$ is constructed as the ratio of the annual citygate price of natural gas in the state to the corresponding residential price of electricity. $RelPrice_{US,t}$ is similarly constructed. All other definitions not noted are as in columns 2, 8 and 9 of Table 2 respectively.

Appendix Table A15: Effect of heating price on mortality: Robustness to excluding control variables

		Dependent variable: Log of all-EWM-causes mortality rate		
		Baseline specification	Winter/non-winter specification using winter	Winter/non-winter specification using HDD
		(1)	(2)	(3)
1	Preferred specification	0.059***	0.073***	0.090***
		[0.017]	[0.019]	[0.032]
2	Exclude housing price index	0.066***	0.074***	0.092***
		[0.016]	[0.019]	[0.032]
3	Exclude unemployment rate	0.058***	0.073***	0.090***
		[0.018]	[0.019]	[0.032]
4	Exclude manufacturing share	0.057***	0.073***	0.090***
		[0.017]	[0.019]	[0.032]
5	Exclude Log(Income) $\times$ Log(RelPrice)	0.041**	0.061***	0.071**
		[0.020]	[0.019]	[0.032]
6	Exclude Share70+ $\times$ Log(RelPrice)	0.059***	0.073***	0.090***
		[0.018]	[0.019]	[0.032]
7	Exclude all pollution and climate controls	0.055***	0.075***	0.094***
		[0.018]	[0.019]	[0.032]
8	Exclude Share70+ $\times$ Log(RelPrice) and Log(Income) $\times$ Log(RelPrice)	0.045**	0.062***	0.071**
		[0.020]	[0.019]	[0.031]
9	Exclude unemployment rate, manufacturing share, and housing price index	0.063***	0.073***	0.092***
		[0.016]	[0.019]	[0.032]
10	Including only basic controls	0.048**	0.065***	0.077**
		[0.021]	[0.019]	[0.031]

Notes: Each cell shows the result from a separate regression, and reports the coefficient on *Heating price proxy* (column 1), *Heating price proxy* $\times$ *Winter* (column 2), or *Heating price proxy* $\times$ *HDD* (column 3). The corresponding standard error, clustered by state, is shown in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Row 1 repeats results from our preferred specifications in columns 2, 8 and 9 of Table 2 respectively. Each row from 2 to 10 shows a change in specification compared to row 1. The change in specification is the exclusion of the control variable(s) indicated in the first column and, where applicable, two-way and three-way interactions that include that variable. *HDD* and variables involving average county HDD in winter months are retained in row 7, column 3. Row 10 retains only county and year-month fixed effects (all columns), all possible two-way interactions between $ShareGas_{j,2000}$, $log(RelPrice_{US,t})$, and *Winter/HDD* (columns 2 and 3), and all possible two- and three-way interactions between $ShareGas_{j,2000}$, $log(RelPrice_{US,t})$, and the average county HDD in winter months (column 3). All other definitions not noted are as in columns 2, 8 and 9 of Table 2 respectively.

B Data appendix

Appendix Table B1 lists the data source for each of our outcome and independent variables. The following sections provide further description of our data sources and the construction of variables used in this paper.

B.1 Mortality rate

B.1.1 Data source

The data source for mortality is Vital Statistics records, specifically restricted-use "mortality files with all county geographical information" obtained from the National Center for Health Statistics (NCHS). These mortality files include a record for every death certificate filed in the United States during the study period. Each record includes a single underlying cause of death, up to twenty additional multiple causes, month of death, and demographic data, including the deceased's age, gender, race, Hispanic origin, education, county of residence and county of death. The definition of the underlying cause of death follows that of the World Health Organization (WHO): the disease or injury which initiated the train of events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury. Causes of death are classified using the Tenth Revision of the International Classification of Disease (ICD-10) during the 2000 to 2010 study period.

We compute mortality rates by county, classifying individuals by their county of residence. We restrict our analyses to the contiguous US throughout the paper because our data source for temperature excludes Hawaii and Alaska. We account for substantial county boundary changes by aggregating counties to a larger stable unit.[15] Specifically, we combine Adams, Broomfield, Boulder, Jefferson, and Weld counties in Colorado; Prince George's and Montgomery in Maryland; Craven and Carteret in North Carolina; Franklin and Gulf in Florida; Bedford and Bedford City in Virginia; Alleghany and Clifton Forge in Virginia; Augusta and Waynesboro in Virginia; Prince William and Manassas Park in Virginia; Southampton and Franklin in Virginia; and York and Newport News in Virginia. We adopt this aggregation throughout the paper.

In addition, when analyzing county-level data, we exclude counties whose population aged 50 and over in 2000 are in the lowest decile of the full sample to reduce noise from mortality rates of counties with small population and missing observations when we use the logarithm of the mortality rate.

B.1.2 Calculating age-adjusted mortality rate

To calculate mortality rates, we use population data from the National Cancer Institutes's Surveillance Epidemiology and End Results (Cancer-SEER) program. These data give yearly county population estimates by age group, sex, race, and Hispanic origin.[16] For 2005, we use the SEER's adjusted set of population estimates that takes into account population shifts due to Hurricanes Katrina and Rita.

We use these population estimates to calculate both crude and age-adjusted mortality rates, expressed per 100,000 population. The crude mortality rate at county-year-month level

[15] Information on substantial county boundary changes was taken from the Census Bureau's website.

[16] We use vintage 2014 population estimates. The data and documentation are available at `https://seer.cancer.gov/popdata/`.

is the total number of deaths in that county in that year-month divided by its population estimate in that year. The age-adjusted mortality rate is a weighted average of the crude mortality rates across age categories, where the shares of each age category in the whole US population are used as weights.[17] We use the age distribution of US population in 2000 (the "US 2000 standard population") published by SEER as weights in the calculation of age-adjusted mortality rates. All mortality rates in the paper are expressed on an annual basis obtained by multiplying the month-level mortality rates by (365/number of day in that month).

B.1.3 Selection of causes of deaths

We use a data-driven approach to select causes of deaths that exhibit significant "excess winter mortality" (EWM), or higher mortality in winter months than in other months.

We use the NCHS's 113 Selected Causes of Death, which represent groupings of detailed ICD-10 codes, as the mutually exclusive set of causes of death. To measure the degree of EWM for each cause, we construct an observation for each month in the 2000 to 2010 period (132 observations) and calculate the total deaths in the US, by cause, in that month. For each cause separately, we run a regression of the number of deaths due to that cause (i.e., as the underlying cause of death) on a *Winter* dummy, which equals 1 for November to March, and year fixed effects. A similar set of regressions is estimated with the logarithm of deaths as the outcome instead of the level. We then select causes whose *Winter* coefficient is in the top quartile among all causes of deaths in both levels and logs (i.e., above 0.12 for logs and 0.02 for levels). We use both levels and logs of mortality because we want to select causes that are both common and have a strong degree of excess winter mortality.

We exclude two causes from the data-driven list of excess winter mortality causes: first, *Accidental exposure to smoke, fire and flames*, since accidental deaths that are not a physiological result of exposure to cold differ from our focus, and second, *All other diseases* (the residual category), since it is difficult to verify the mechanism for this "cause." Appendix Table A3 reports *Winter* coefficients in levels and logs and average monthly crude mortality rate for each of the selected causes. The final selected list includes the following fourteen causes of death, with their ICD-10 codes in brackets. These causes can be further grouped into four broader cause groups, and generally match the causes highlighted in the health literature as exacerbated by cold.[18]

- **Group A:** Non-viral, non-respiratory infections

[17] We use the following 19 age categories: under 1 year, 1-4 years, 5-9 years, ..., 80-84 years, and 85 years and over.

[18] In cold temperature, blood vessels constrict to conserve heat and maintain body temperature, causing higher cardiac workload and blood pressure (Castellani and Young 2016; Keatinge et al. 1984). These factors along with changes in blood chemistry (including increased levels of fibrinogen, cholesterol, and platelet aggregation) increase the risk of adverse cardiovascular events such as strokes, myocardial infarctions, and pulmonary embolisms (Crawford et al. 2003; Liddell and Morris 2010; Woodhouse et al. 1994). Exposure to cold temperature is also associated with increased incidence and severity of respiratory tract infections and exacerbation of chronic respiratory diseases (Donaldson and Wedzicha 2014; Mourtzoukou and Falagas 2007). The mechanisms linking cold weather to these respiratory problems include increased broncho-constriction and compromised local respiratory defenses due to the inhalation of cold air (Donaldson and Wedzicha 2014; Eccles 2002).

  - Septicemia (A40-A41)

- **Group G:** Neurological diseases
  - Parkinson's disease (G20-G21)
  - Alzheimer's disease (G30)

- **Group I:** Circulatory system diseases
  - Acute myocardial infarction (I21-I22)
  - All other forms of chronic ischemic heart disease (I20, I25.1-I25.9)
  - Heart failure (I50)
  - Cerebrovascular diseases (I60-I69)
  - Atherosclerosis (I70)

- **Group J:** Respiratory system diseases
  - Influenza (J09-J11)
  - Pneumonia (J12-J18)
  - Emphysema (J43)
  - Other chronic lower respiratory diseases (J44, J47)
  - Pneumonitis due to solids and liquids (J69)
  - Other diseases of respiratory system (J00-J06, J30-J39, J67, J70-J98)

B.2 Life-years lost

Life-years lost (reported in Table C1) are calculated by combining our estimated mortality impacts with life expectancy estimates from the 2000 United States life tables published by the National Center for Health Statistics. Using the life tables, we first compute the residual life estimate for each of the three age groups (under 65, 65-74 and 75 and over) as the weighted average of residual life estimate for each single age in the age group, where the weights are the population proportion of each age in 2000.[19] The winter life-years lost impact for each age group is the product of the winter mortality impact and the residual life estimate for the age group and is expressed in terms of annualized life-years lost per 100,000 population. We then take the weighted average across the three age groups, where the weights are the population proportion of the age group, to obtain the winter life-years lost impacts reported in Table C1.

To compare the value of life-years lost with the cost of heating incurred by households, we first convert the winter life-years lost effect per 100,000 *people* into annual life-years per *household*, using data on the total US population, proportion of the US population living in households, and number of households in the US during our sample period 2000-2010.

[19]The residual life estimates for the three age categories are 48.5 years (under 65), 14.9 years (age 65-74), and 8.5 years (age 75 and over).

B.3 Home energy price and usage

All energy prices and consumption data come from monthly series published by the US Energy Information Administration (EIA), available at the state and national level. The data are based on samples of firms supplying natural gas or electricity to residential consumers, and include some processing by EIA to account for non-response.[20]

The raw data express quantities in kilowatt-hours for electricity and cubic feet for natural gas. To allow comparison between energy types, we convert these quantities to British Thermal Units (BTU). The conversion is straightforward for electricity. For natural gas, we apply estimates of the heat content of natural gas delivered to residential consumers for each state and year using the company-level data available in EIA's Natural Gas Annual Respondent Query System. For these estimates, we drop firms reporting heat content values of 0 or above 2,500 BTU per cubic feet, and weight the reported heat content for each firm by the volume of gas supplied to residential consumers.[21] We also apply two manual edits. First, five state-year observations are missing residential consumer heat content data for all firms; we use the all-consumers heat content for these five observations. Second, the dominant firm in Arkansas is missing heat content data for 2001; we use the average of its report in 2000 and 2002 instead.

Lastly, to aid interpretation of monetary units, we deflate all prices in this paper—including the prices of natural gas and of electricity—to 2016 prices using the Bureau of Labor Statistics's (BLS) Consumer Price Index (CPI-U).

B.4 Home energy bills

For data on energy bills, we use Census 2000 5-Percent Public Use Microdata Sample (PUMS) files combined with the 2005 to 2010 American Community Survey (ACS) PUMS files (Ruggles et al. 2022).[22] The Census/ACS data are available on an annual basis, and the finest geographic identifier is the Public Use Microdata Area (PUMA). We aggregate the microdata to obtain mean monthly energy bill for each PUMA for the year 2000 and 2005–2010. We then crosswalk from PUMA-level to county-level to facilitate discussion.[23]

The relevant question in the Census 2000 is "What are the annual costs of utilities and fuels for this house, apartment, or mobile home?", broken down into different types of utilities and fuels. In the ACS, households are asked how much these bills cost them last month (for electricity and gas) and last 12 months (for other fuels). We exclude households whose energy bills are included in their rent or condominium fees.

B.5 Home heating sources ($\boldsymbol{ShareGas_{jt}}$, $\boldsymbol{ShareGas_{j,2000}}$)

$ShareGas_{jt}$, the proportion of occupied housing units that indicate that gas is their main heating source in each year, is computed using state- and county-level aggregate data available in the 2000 Decennial Census Summary Files and the 2005–2009, 2006–2010, 2007–

[20] Response to the survey is required by law, and hence such non-response should not be a large problem.

[21] Heat contents typically range between 900 and 1,200 BTU per cubic feet.

[22] The data are self-reported but the reporting error is not likely to be correlated with the price of heating. As such, these errors will likely reduce precision but not affect the interpretation of our estimated effect.

[23] We use the crosswalk from the Missouri Census Data Center, available at `https://mcdc.missouri.edu/applications/geocorr.html`.

2011, and 2008–2012 ACS 5-year estimates.[24] The ACS estimates are matched to the mid-point of the 5-year range. This procedure yields state- and county-year $ShareGas_{jt}$ for 2000, 2007, 2008, 2009, 2010. For the years 2001 to 2006 without Census or ACS data, since $ShareGas_{jt}$ is highly correlated over time, we obtain $ShareGas_{jt}$ by interpolation.

The relevant Census or ACS question from which $ShareGas$ is derived is "Which fuel is used most for heating this house, apartment, or mobile home?" We include both utility gas from underground pipes serving the neighborhoods and bottled, tank or LP gas in the definition of "gas".[25] Our instrument is constructed using $ShareGas_{j,2000}$, which is the value of $ShareGas_{jt}$ in the year 2000.

B.6 Relative price of gas to electricity ($\boldsymbol{RelPrice_{jt}}$, $\boldsymbol{RelPrice_{US,t}}$)

Two other key variables in our analysis are $\log(RelPrice_{jt})$ and $\log(RelPrice_{US,t})$, the log of the relative price of natural gas to electricity in the state-month and US-month, respectively. We use the same data described in Section B.3 for this. The electricity price for the denominator was described described previously in that section.

One candidate for the natural gas price is the monthly price of natural gas delivered to residential consumers, which is computed by dividing the reported revenue of local distribution utilities by the associated sales volume. The relevant survey question that EIA uses defines revenue as "gross revenues including any and all demand charges, commodity charges, taxes, surcharges, adjustments or other charges billed for gas delivered"; consequently, fixed charges that utilities frequently include (e.g. basic monthly customer charges that do not depend on volumes) are included. However, we expect consumers to respond to the variable (i.e., usage-dependent) component of prices, not the fixed charge component. In the data, since the fixed charges are averaged over a smaller volume in summer, the residential price spikes in summer (Appendix Figure B1).

Because of this, we use the monthly price of natural gas at the citygate instead. The citygate price is the price faced by local distribution utilities (companies that sell gas to residential consumers); hence it captures variation due to natural gas prices and excludes fixed charges to residents. In addition, utilities are required by federal law to price gas on a cost-recovery basis.[26] This means that absent forecast errors by the utilities, citygate prices should capture the variation in gas price perfectly. With forecast errors, utilities are legally required to return unexpected profits or losses made on natural gas to consumers by adjusting the future months' prices downwards or upwards. Citygate prices are not available for electricity.

When examining home energy bills and in robustness-check specifications, we use the annual versions of the price variables to match the timing of the outcome variable. These are based on a separate survey of the universe of firms in the US, but are otherwise identical to the monthly versions.

[24] For the 2000 Census, we use Table H40 of Summary Files 3. For the ACS, we use data frome National Historical Geographic Information System (NHGIS; Manson et al. 2017).

[25] Other energy sources used for heating include fuel oil, kerosene, coal, coke, wood, and solar.

[26] Note that they may still charge a markup on distribution of gas, which is more difficult for the state to monitor.

B.7 Heating degree-days

To compute the number of heating degree-days (HDD), we use daily gridded temperature data for the contiguous US (4 kilometers by 4 kilometers resolution) from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) data developed and maintained by the PRISM Climate Group at Oregon State University (PRISM Climate Group 2004).[27] The PRISM data incorporate the current knowledge of US spatial climate patterns, including elevation and prevailing wind patterns, and are the official spatial climate datasets of the US Department of Agriculture (Daly et al. 2008).

To obtain HDD for each county-month, we first compute the geographic average daily mean temperature of each Census 2000 block group. For each block group, we take a simple average of all grid points within, or on the boundary of, the block group. We then compute HDD for each month for each block group, based on

$$HDD_{it} = \sum_{x=1}^{T(t)} \max \{threshold - tmean_{ix}, 0\} \tag{5}$$

where HDD_{it} is the HDD of block group i in month t, $threshold$ is a temperature threshold (set at 65°F, following convention), $tmean_{ix}$ is the mean temperature of block group i on day x, and $T(t)$ is the number of days in month t. Next, we compute each county's HDD for the month by taking the average of the block groups within the county, weighted by the population in Census 2000. Finally, we scale HDD to a 30-day month, and divide by 1,000, to yield an average monthly measure of coldness. Block group geographic and population data come from the NHGIS.

B.8 Household income and population share age 70+

Data for county and state household income and fraction of people age 70 and above are from the 2000 Decennial Census of Population and Housing Summary Files. Both variables are derived from the Census using the same approach as described above for *ShareGas*.

The Summary Files do not report the 25th and 75th percentiles of household income at the county level. Hence, these variables are constructed using tract-level data on the number of households in 16 income bins, available in the Summary Files (we use the NHGIS version). Specifically, we interpolate the proportion of households in the income bins to obtain the 25th and 75th percentiles of household income at the tract level, and then aggregate these variables up to the county or state level, weighted by the number of households.

B.9 House price index

State house price index used in the paper is the quarterly seasonally-adjusted purchase-only house price index, available from the Federal Housing Finance Agency (FHFA).

B.10 Unemployment rate

We use the Bureau of Labor Statistics' county-month unemployment rate as a control variable. A few county-level observations are missing due to Hurricane Katrina; we

[27] The specific dataset used is version D1 of the AN81d dataset, retrieved February 2017, from `http://prism.oregonstate.edu`.

use the state unemployment rate for these observations, and include a dummy for affected observations in regressions.

B.11 Manufacturing share of the economy

We use the Bureau of Economic Analysis's state-quarter personal income data when controlling for manufacturing share of total employee compensation (meant to proxy for share of the economy). A few observations (fewer than 0.5%) are missing; we impute these observations by interpolation. Quarterly data are then matched to the appropriate time period.

B.12 Absolute humidity

We use block group-month level temperature and dewpoint temperature in the PRISM data to compute absolute humidity. Absolute humidity (in grams per cubic meters) is computed using the psychrometric formulas in Snyder and Melo-Abreu (2005, Appendix 3)

$$AbsoluteHumidity = \frac{2165 \times VaporPressure_{dew}}{Temperature + 273.16} \tag{6}$$

$$VaporPressure_{dew} = 0.6108 \exp\left(\frac{17.27 \times Temperature_{dew}}{Temperature_{dew} + 237.3}\right) \tag{7}$$

where $Temperature$ is the temperature in degree Celsius, and $VaporPressure_{dew}$ is the vapor pressure in kilopascals computed at the dewpoint temperature $Temperature_{dew}$ in degree Celsius. We then aggregate to the county-month level, weighting by population in 2000.

B.13 Air pollution data

The data source for air pollution is daily station-level data from the US Environmental Protection Agency (EPA) Air Quality System (AQS).[28] The AQS data contain daily air quality indices (AQIs) for carbon monoxide, nitrogen dioxide, ozone, particulate matter (2.5 and 10), and sulfur dioxide; data for some pollutants for some stations are missing. We construct monthly AQIs for each geographic unit of analysis, and then aggregate to the appropriate time period.

We use a mix of procedures to construct monthly AQIs. For the first procedure, we compute each pollutant's AQI at the Census 2000 block group level, and then aggregate to the county or state level, weighting by population. We compute the AQI for each block group as the average of all AQI measurements taken within a month at all stations within 100 kilometers, weighted by the inverse of the squared distance to the station. County AQI is then the population-weighted average of block group AQIs. Block group and population data come from the NHGIS.

The above procedure (setting a distance threshold and computing the AQI) is standard in the literature using stations data, but it produces many missing observations at the county level. We patch missing data using a second procedure. Specifically, if a county has more than 50 percent of its population not assigned a pollutant AQI value in any month, we use a second

[28]The EPA provides several ways of accessing the data. We use the pre-generated data files, accessed February 2017.

procedure to compute its AQI values for all months. For these counties, for each month, we compute AQI based on the nearest five stations with available measurements, weighted by the inverse of the squared distance between the station and the county centroid. This guarantees that all counties have a pollution AQI measure for all months in consideration.

Appendix Table B2 shows, for each pollutant, the breakdown of the procedure used to compute AQI in the sample of counties used in our analysis of mortality data.

B.14 Independent variables used in heterogeneity analysis

Variables used in analysis of heterogeneity in the effect of heating price on mortality are from the following sources:

- **Poverty rate:** Data on proportion of households in the county with income below 150% of the poverty level is from the 2000 Decennial Census of Population and Housing Summary Files.

- **Education:** Data on the deceased's education level is provided in the mortality files. We drop deaths that occur before the age of 25, with censored education level, for this analysis. To compute age-adjusted mortality rates by education level, we use Census/ACS population data, since the SEER data does not contain a breakdown by education level. We interpolate proportions for the years in which no population data exist (2001 to 2004). Note that education information is missing for about 5.4 percent of individuals aged 25 and over in our dataset. This, along with some misreporting on the death certificate (Rostron et al. 2010), adds noise to our heterogeneity estimates by education.

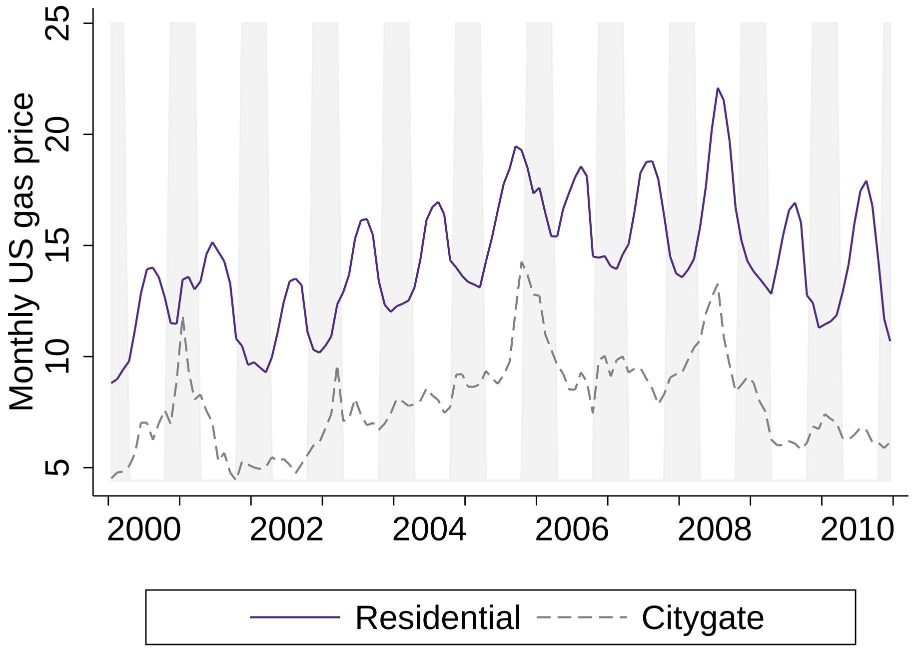
- **Sex:** Data on the deceased's sex is provided in the mortality files.

- **Race:** Data on the deceased's race is provided in the mortality files. Following Schwandt et al. (2021), we include both non-Hispanic and Hispanic persons under "Black" or "White", and exclude American Indian, Alaska Native, Asian, Native Hawaiian, and Other Pacific Islander persons from either category.

B.15 LIHEAP data

As a robustness check, we use data on Low Income Home Energy Assistance Program (LIHEAP) spending from the US Department of Health and Human Services' LIHEAP Data Warehouse. The data are based on mandatory reports from states for each fiscal year, and are available at the state-fiscal year level starting in fiscal year 2001 (i.e. since October 2000). For the nine months in our sample without LIHEAP data, we impute an arbitrary value for LIHEAP per capita and include a dummy for affected observations in the regressions.

Appendix Figure B1: National price of natural gas over time



Notes: Price in dollars per million BTU. Gray regions are winter months (November–March).

Appendix Table B1: Data sources

Data	Data source	Geographic identifier	Temporal identifier
Dependent variables			
Mortality rate	Vital Statistics Mortality Files	County	Month
Average home energy price	Energy Information Administration (EIA)	State	Month
Home energy usage	Energy Information Administration	State	Month
Home energy bill	Census; American Community Survey (ACS)	PUMA	Year
Independent variables			
Home heating energy type	Census	Census tract	Year
Energy prices	Energy Information Administration	State	Month
Temperature	PRISM	Grid point[a]	Day
Median household income	Census	Census tract	Year
Fraction of people aged 70 & above	Census	Census tract	Year
House price index	Federal Housing Finance Agency	State	Quarter
Absolute humidity	PRISM	Grid point[a]	Month
Air pollution	Environ. Protection Agency Air Quality System	Pollution monitor	Day
Unemployment rate	Bureau of Labor Statistics	County	Month
Manufacturing share of economy	Bureau of Economic Analysis	State	Quarter
LIHEAP assistance funds	Department of Health and Human Services	State	Fiscal year

[a] 4 km by 4 km resolution

Appendix Table B2: Frequency of the two interpolation procedures used for calculating AQIs

	CO	NO_2	O_3	$PM_{2.5}$	PM_{10}	SO_2
Based on distance threshold	1,177	1,048	1,096	2,231	1,762	1,512
Based on nearest 5 stations	1,616	1,745	1,697	562	1,031	1,281
Total counties	2,793	2,793	2,793	2,793	2,793	2,793

C Additional details and results

This appendix discusses additional details on our specification and results.

C.1 Winter/non-winter specification with HDD as additional dimension

Equation (4) in the main text shows our winter/non-winter specification with Winter as the additional dimension. As mentioned, since some winters or particular months in winter are colder than other, we also use HDD to define the additional dimension. Specifically, we estimate the following specification:

$$
\begin{aligned}
log(m)_{jt} \quad = \quad & \alpha + \lambda_1 \,\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt}) \times HDD_{jt} \\
& + \lambda_2 \,\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt}) + \lambda_3 \,\text{ShareGas}_{j,2000} \times HDD_{jt} \\
& + \lambda_4 \log(\text{RelPrice}_{\text{US},t}) \times HDD_{jt} + \theta_1 Z_j \times \log(\text{RelPrice}_{\text{US},t}) \times HDD_{jt} \\
& + \theta_2 Z_j \times \log(\text{RelPrice}_{\text{US},t}) + \theta_3 Z_j \times HDD_{jt} \\
& + \theta_4 \,\text{ShareGas}_{j,2000} \times \log(\text{RelPrice}_{\text{US},t}) \times \overline{HDD}_j \\
& + \theta_5 \log(\text{RelPrice}_{\text{US},t}) \times \overline{HDD}_j + \theta_6 Z_j \times \log(\text{RelPrice}_{\text{US},t}) \times \overline{HDD}_j \\
& + \gamma_j + \tau_t + \delta X_{jt} + \epsilon_{jt}
\end{aligned} \tag{8}
$$

In this specification, we control for the county's average HDD in winter, $\overline{\text{HDD}}_j$, in parallel to HDD_{jt} to adjust for systematic differences (e.g., demographics) between colder regions and warmer ones. Note that since we control for average HDD in winter, the coefficient on the heating price proxy, $\text{ShareGas}_{jt} \times \log(\text{RelPrice}_{jt})$, is the effect in warm months for the theoretical county with exactly zero HDD in all winter months. This coefficient relies heavily on an extrapolation because there are few observations with winter HDD of zero — the median and 5th percentile of winter average HDD (across counties) are 0.78 and 0.31 respectively. Because of this, the coefficient on the heating price proxy is not interpretable.

C.2 Robustness checks and threats to validity

This appendix section assesses whether our results are robust to varying our specification, and investigates potential threats to the validity of the research design. Appendix Table A14 reports these robustness checks for the baseline and winter/non-winter specification estimates; our preferred specification is reproduced as the first row of the table.

The first two robustness checks vary which months we define as winter. Our main specification uses November to March, based on the monthly pattern of average temperature in the US. We find similar results if we use December to March, like the Europe-focused literature on excess winter mortality, or as December to February, the three coldest months of the year.

Winter price spikes might affect mortality in subsequent non-winter months, and potentially bias our winter/non-winter specification coefficients. To investigate this, our next robustness check specification uses the previous non-winter months as a comparison group, and finds little difference in our results (row 3).

Next, we vary how we construct our energy price variable. Our main specification uses the citygate natural gas price, which more closely reflects the marginal price that residential consumers face, as opposed to the residential price. For electricity prices, no citygate price data exist, so we use the residential price. Our results are robust to using the residential price for natural gas too. We also show that our results are robust to using the annual instead of monthly natural gas price.

We construct ShareGas as the proportion of all households that use gas for heating and then focus on gas and electricity prices. While these are the two most common energy sources for heating, there are other sources too. We show that the results are not sensitive to this simplification. First, we use an alternative definition of ShareGas, which is the share among households that use either gas or electricity for heating (row 7). We also show the results excluding states in which the share of households that use gas or electricity for heating is less than 75%, such as those in New England where fuel oil is a common heating source (row 8).

One potential concern is that shale gas production in a community itself could affect mortality and might also make it more likely that natural gas is the energy source used locally for heating. To address this concern, we estimate the results excluding all states that produce shale gas. Here too, the results are very similar to our main results.

The main specification includes control variables to address the fact that the Great Recession overlaps with the study period. We also go further, as a robustness check, and drop the Great Recession period, as defined by the NBER Business Cycle Dating Committee (row 10). Another type of concern is that the government LIHEAP program might respond to or be spuriously correlated with heating prices. We thus control for the state's per capita spending on LIHEAP in row 11.[29] Next, our main specification focuses on particulate matter and nitrogen dioxide as these are the pollutants correlated with mortality; in row 12, we include all the air pollution variables as controls, instead of just those that are most linked to mortality.

Additionally, our results are robust to using a more disciplined approach to selecting controls based on the double-selection post-LASSO method of Belloni et al. (2014). Specifically, we start with a rich set of 45 potential controls spanning demographic, health care, temperature, pollution, and social determinants of health (14 of which are our baseline control variables).[30] From this set, the double-selection procedure selects by LASSO a set of

[29]LIHEAP state-year spending data are from the Department of Health and Human Services.

[30]In addition to the controls in the main specification, this set includes additional climate variables constructed using the PRISM data (monthly average temperature, the averages (over days) of the minimum and maximum temperatures, average dewpoint temperature, average precipitation, and the average maximum and minimum vapor pressure deficits), pollution variables from the AQS data (the AQIs of carbon monoxide, ozone, and sulfur dioxide), health-related variables from the Social Determinant of Health dataset published by the Agency for Healthcare Research and Quality (log of number of long-term hospitals, log of number of short-term general hospital beds, log of number of nursing home beds, log number of Medicare eligibles in the county, proportion of 20+ population that is obese and proportion diagnosed with diabetes), LIHEAP control (log of total LIHEAP assistance funds per capita in the state-fiscal year), along with state-level number of doctors per capita from Barreca et al. 2016, and demographic and housing characteristics from the census (proportion of population that is female, proportion of population that is foreign-born, median age, proportion of population White, proportion of civilian population consisting of veterans, average household size, proportion of households that received food stamps/SNAP, proportion of households that are below 150% of the federal poverty line, proportion of 25+ population with less than high school education, propor-

variables that are useful for predicting log EWM mortality, a set of variables useful for predicting the local heating price proxy, and a set of variables useful for predicting our instrument. We then estimate the effect of heating price on mortality with the union of these three sets of selected variables as included controls. Applied to our baseline specification, this procedure selects 6 variables: the interactions of $log(RelPrice_{US,t})$ with the log of county median household income in 1999, the percent of housing units without fuel and the percent of housing units that are mobile homes in 2000, NO_2 squared, absolute humidity, and dew point. Column 1 of Row 13 reports our IV estimate controlling for these LASSO selected variables. Our winter/non-winter specification estimates in columns 2 and 3 control for these selected variables plus additional two- and three-way interactions analogous to our regressions in columns 8 and 9 of Table 2. The estimates are qualitatively similar to our baseline results.

Figure 1 suggests that much of the variation in ShareGas is between states. For this reason, our main specification clusters standard errors at the state level. We can instead estimate the regressions at the state-month level, using only between-state variation. As shown in row 14, the results are similar. To check if the results are driven by variation at an even larger geographic scope, we also estimate the state-month level regressions controlling for Census division fixed effects interacted with $\log(\text{RelPrice}_{\text{US},t})$ (row 15). The baseline specification coefficient remains very similar (with a larger standard error), while the winter/non-winter specification coefficients are no longer significant but show a broadly similar pattern. Finally, row 16 uses annual variation in RelPrice in state-month-level regressions and obtains similar results.

Appendix Table A15 reports another set of robustness checks in which we remove different control variables from our main specification. Just as our main results are robust to adding additional control variables, they are also robust to removing control variables.

We view the results in this appendix as supporting the validity of the finding that a higher heating price causes an increase in winter mortality.

C.3 Welfare computation: effect size expressed in life-years lost

Our main analysis investigates the impact of heating prices on the extensive margin of mortality, but it is also valuable to understand the intensive margin, or how many years of life are lost. For this, we start with heterogeneity in the effect by age groups, reported in Appendix Table C1. Not surprisingly, we find that most of the mortality averted when heating prices are lower is among older age groups.

We report in the last row of Appendix Table C1 the impact of heating prices on life-years lost, which takes into account the remaining life expectancy of those who die (or whose deaths are averted). Combining our estimated mortality effects with residual life estimates from the life tables for 2000 published by the National Center for Health Statistics, we find that a 1% increase in the price of heating causes around 7 annualized life-years lost per 100,000 people during winter, or equivalently 3 annual life-years lost per 100,000 people. (The data appendix provides further details on this calculation.)

Using $369,000 as the value of a statistical life-year (Kniesner and Viscusi 2019), our

tion of persons in institutionalized group quarters, proportion of housing units that are rented, proportion of housing units that are mobile homes, and proportion of occupied housing units without fuel).

estimates imply that a 1% increase in the price of heating leads to \$27.27 (in 2016 dollars) of life-years value lost per household. We can compare this cost to the amount of money needed to avert those deaths. This sheds light on whether mortality risk and money are being traded off in a way that seems socially optimal.[31]

We calculate the amount of money needed to avert the heating-price induced deaths two ways. First, assume that the mortality occurs exclusively through the heating consumption channel — from cutbacks in heating use — rather than the non-heating consumption channel. As discussed in Section 4.1, the increase in energy bills when the price of heating rises is less than one-for-one, because households cut back on heating. Our estimates of how heating prices affect energy bills in Table 1 imply that a household reduces spending on heating by \$1.98 annually in response to a 1% price increase, compared to if they had not cut back on heating.[32]

Thus, the life-years cost is more than 13 times as large as the extra outlay on heating needed to avert the mortality, which is consistent with households either facing credit constraints or other frictions or not optimizing. The benefit of compensating households for increases in heating bills is therefore significantly larger than the cost. This suggests that expanding safety-net programs such as LIHEAP would be welfare-enhancing. The increase in benefits could be indexed to the weather or the price, offering households more compensation when the temperature is abnormally cold or the price especially high.

A second way of calculating the outlay needed to avert the mortality is to assume that one would need to offset all of the cutbacks in other spending too (non-heating consumption channel) in addition to the heating consumption channel. Then, the cost is simply 1% multiplied by the starting-point heating expenses. This amount is \$8.06, which is still much smaller than the monetary value of the lives lost. Importantly, the \$8.06 value is an upper bound on the outlays needed because some cutbacks in spending that households make in response to higher heating bills (e.g., restaurant dining, cigarettes) do not increase their mortality risk.

Our empirical results are suggestive that the heating consumption channel is the main one, so the first of the two calculations seems like the more appropriate one. It implies that the lives lost due to cutbacks in heating when heating prices increase are more valuable than the additional spending required to avert those deaths.

Our findings have implications for several types of policies that can reduce households' heating costs. For example, they help quantify a potentially important benefit—averted deaths—of the federal Low Income Home Energy Assistance Program (LIHEAP), which assists low-income households with their energy bills, and state energy price subsidy programs (such as the California Alternate Rates for Energy).[33] The computations above suggests that increasing compensation to households so that they avoid such large cutbacks would be cost-

[31] If we assume that households are in fact optimizing, then we could instead use this analysis to calculate a new revealed-preference value of a statistical life year, which would be \$26,800. However, the assumption that low-income households are not credit constrained would be a tenuous one.

[32] To calculate spending on heating, given that our regressions estimate impacts on the total energy bill, we use the fact that 29% of a household's energy spending is on heating (RECS 2009).

[33] Hahn and Metcalfe (2021) evaluate the welfare impacts of the California subsidy program that arise through economic redistribution and environmental costs; their analysis does not directly assess its health effects.

effective. LIHEAP payments do not typically increase when the price of heating increases, which leaves the poor uninsured against this risk, so the findings also point to potential design improvements for LIHEAP.[34] The results are also relevant for cost-benefit analysis of weatherization programs that reduce households' need for heating.

C.4 Welfare comparison of value of averted life lost

Based on a spatial equilibrium model, Bartik et al. (2019) estimate a net willingness-to-pay for fracking of \$2,500 per household annually among households living in shale play counties. This estimate takes into account local factors like increased economic activity and possibly negative changes in amenities due to fracking. In their replication kit, they also report that there are 15.9 million households living in 16 main shale plays, which implies that the total US local-level welfare gain is \$40 billion a year.

Next, our results imply that the decline in the price of natural gas relative to electricity between 2005 to 2010 averted 12,5000 deaths per year. Using the value of a statistical life (VSL) of \$10 million, this maps to a gain of \$125 billion a year. Alternatively, since most of the mortality averted is among older individuals, we can quantify this benefit in terms of the value of life years lost averted. Using a similar calculation as that discussed in Appendix C.3, this represents 279,000 life-years lost per year which maps to \$103 billion using the value of a statistical life year (VSLY) of \$369,000 in 2016 dollars (Kniesner and Viscusi 2019). Shale gas production is estimated to explain 83% of the decline in natural gas price during this period (Hausman and Kellogg 2015). This means that the national-level benefit of fracking in the form of averted deaths is more than twice as large as the local-level net welfare gain, and should not be ignored when considering the effects of shale production of natural gas.

Finally, again using the VSL of \$10 million (or VSLY of \$369,000), our results map to \$4.2 billion in terms of averted deaths (or \$3.2 billion in terms of averted life-years lost) for each percent decrease in the heating price. Davis and Kilian (2011) estimate that the US residential natural gas price ceilings of 1954–89 imposed a welfare cost of \$3.6 billion annually in 2000 dollars. Comparing the two estimates, and adjusting for inflation and differences in the number of households within the two periods, the health benefit of a price decrease of 2.25 percent would have outweighed the allocative costs.

[34] On average between 2001 and 2010, 4.5% of US households received LIHEAP heating assistance per year, which is 23% of households below 150% of the poverty line. LIHEAP pays eligible households a preset amount each year based on income and household size, and, depending on the state, also fuel type or the last year's utility bills. Arizona is, to our knowledge, the only state that varies the amount based on contemporaneous bills or prices (LIHEAP Clearinghouse 2010).

Appendix Table C1: Heterogeneous effects on mortality by age groups

	Dependent variable: Mortality rate	
	All causes	All EWM causes
	(1)	(2)
Heating price proxy	10.5	9.72**
	[6.50]	[4.37]
Heating price proxy × 65–74	112.8	94.9*
	[70.8]	[56.4]
Heating price proxy × 75+	254.8*	262.0**
	[133.1]	[121.5]
Observations	460,020	460,020
Mean mortality rate	3835.3	2663.5
Implied mortality effect for 65–74 population	123.32*	104.62*
	[70.54]	[56.52]
Implied mortality effect for 75+ population	265.30*	271.71**
	[133.95]	[121.41]
Implied effect on life-years (per 100,000 population)	6.99	6.50

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises county-year-months-age group for winter months. Mortality rates are age-adjusted mortality rates expressed as annual deaths per 100,000 population; see data appendix for further details. *Heating price proxy* is $ShareGas_{jt} \times Log(RelPrice_{jt})$, where $ShareGas_{jt}$ is the county-year proportion of occupied housing units with natural gas as their main heating source, and $RelPrice_{jt}$ is the log of the ratio of the state-month citygate price of natural gas, averaged over the three- and four-month lag, to the corresponding residential price of electricity. *65–74* is an indicator variable that equals one for the 65 to 74 population. *75+* is an indicator variable that equals one for the 75 and over population. *Heating price proxy* and its interaction with *Trait* are instrumented using $ShareGas_{j,2000} \times Log(RelPrice_{US,t})$ and its interaction with *Trait*. All columns include all fixed effects and control variables from column 2 of Table 2, the main effect for each age group indicator, and the interaction of each fixed effect or control variable with each age group indicator. *Implied mortality effect* for each age group is the coefficient on the heating price proxy plus the coefficient on the analogous interaction term. *Implied effect on life-years* is the implied effect of a 1% increase in heating price on life-years per 100,000 population. It is the weighted sum of the effect on life-years across the 3 age groups, each calculated as the product of the mortality effect and the average residual life estimate for the age group based on the 2000 National Vital Statistics Reports, where the weights are the proportion of the age group in the US population in 2000. See data appendix for further details.