Surviving the Winter: Inexpensive Heating Reduces Mortality*

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

[1]Northwestern University, Department of Economics
[2]National Bureau of Economic Research

PRELIMINARY AND INCOMPLETE DRAFT
November 12, 2017

Abstract

This paper examines how the price of heating affects mortality in the US. When the price is higher, households will use less heating, and exposure to cold has been linked to cardiovascular failure, respiratory infections, and other health problems, and is a hypothesized contributor to the pattern of "excess winter mortality." Heating prices could also affect health if high energy bills lead to cutbacks in food or health care spending. Our empirical approach combines spatial variation across the US in the energy source used for home heating and temporal variation in the national prices of different energy sources; whether natural gas or electricity is used for heating varies significantly across counties, and the relative price of natural gas to electricity varies over time, notably due to the boom in shale production of natural gas during our 2000 to 2010 study period. Using microdata on all deaths in the US, we find that a lower heating price reduces winter mortality. The effect is larger for males and is due mostly to cardiovascular and respiratory causes of death.

*Chirakijja: janjalachirakijja2012@u.northwestern.edu; Jayachandran: seema@northwestern.edu; Ong: pinchuanong2014@u.northwestern.edu. We thank Sachet Bangia, Alejandro Favela, and Caitlin Rowe for outstanding research assistance. All errors are our own.

1 Introduction

Many families spend a large share of their income on home energy bills. For example, 17% of American households spend more than 10% of their income on home energy bills. Heating is the single largest contributor to annual home energy consumption in the US (despite being used for only part of the year) (RECS 2009).

Heating costs present households with a difficult trade-off: They have to either keep their home uncomfortably cold to save on heating or forgo other spending to afford their high heating bill. How acute this dilemma is depends on how expensive home heating is. Through both a substitution and an income effect, a higher price of heating could be detrimental to health. First, using less heating (substitution effect) means exposure to lower ambient temperature, which has been linked to cardiovascular problems, respiratory infections, and other health problems. Second, if families do not cut back usage one-for-one when the price rises, then their energy bills will increase. This can lead to cutbacks in other expenditures that affect health, such as food and health care (income effect).

This paper estimates the causal effect of heating prices on mortality in the US. A large literature has documented that mortality, especially of the elderly, peaks in winter and that cold weather is associated with higher mortality. Our contribution is to examine whether high home heating costs exacerbate the pattern of "excess winter mortality."

The empirical design uses spatial variation across the US in the energy source used for home heating. Natural gas and electricity are used by 58% and 30% of American households, respectively. Importantly, there is a large amount of geographic variation across US counties in whether an area relies on natural gas versus electricity. We combine this spatial variation with temporal variation in the national prices of natural gas and electricity. The ratio of the natural gas to electricity price varied considerably over the 2000 to 2010 study period, most notably due to the boom in shale production of natural gas (commonly known as "fracking" because of the hydraulic fracturing method used to extract shale gas). We leverage the fact that households in areas that rely more on natural gas for heating experienced a decline in their home heating price as a result of the fracking boom, relative to households in areas more reliant on electricity.

Using this approach, we find that lower heating prices reduce winter mortality, where for the purposes of this paper, we define "winter" as November to February. The effect is driven mostly by male mortality and cardiovascular and respiratory causes. As our main outcome measure, we zero in on age-adjusted mortality from causes that exhibit a strong pattern of excess winter mortality.[1] These causes account for about half of total mortality. We also find a significant effect using all-cause mortality. Our results are robust to several checks on the specification, such as dropping states in which shale production of natural gas occurs.

Our paper contributes to the literature on the effects of cold weather and wintertime on mortality (The Eurowinter Group 1997; Analitis et al. 2008; Deschenes and Moretti 2009), morbidity (Ye et al. 2012), and nutrition (Bhattacharya et al. 2003; Beatty, Blow, and Crossley 2014). To our knowledge, ours is the first study to estimate the causal effect of heating prices — a plausibly important and policy-relevant mediating factor — on mortality or, more generally, on health. Previous work has found that the spike in winter mortality is stronger among those living in older housing, which tends to be less well-insulated, which is suggestive but not dispositive that indoor home temperature is a mediating factor (Wilkinson et al. 2007). The studies closest to ours are those that examine how home weatherization affects health; some studies report reductions in morbidity, and others find null results (Critchley et al., 2007; Howden & Chapman, 2007; Green & Gilbertson, 2008; El Ansari & El-Silimy, 2008). Most of these studies analyze small samples and thus lack the statistical power to examine mortality or objectively measured health outcomes. Another related literature documents a positive association between heating subsidies for low-income families and health; the studies are not designed to establish causal relationships, however (Frank et al. 2006).

Our paper also contributes to the literature on the consequences of the boom in shale production of natural gas. Most of the previous literature in economics has focused on local economic effects such as job creation and higher wages (Feyrer, Mansur, and Sacerdote 2017), as well as possible negative consequences such as the influx of cash increasing crime (DeLeire, Eliason, and Timmins 2014; Bartik et al. 2017). We, instead, examine health effects that occur via impacts on energy prices. Note that fracking likely affects health through other channels too. Shale gas often displaces coal in electricity generation, lowering emissions of particulate matter and other pollutants (Jackson et al., 2014). However, there are also potentially large local environmental health costs of fracking, for example due to chemical

[1] We use a data-driven approach that examines seasonal patterns at the national level for each of the National Center for Health Statistics' 113 selected causes of death and selects those that are in the top quartile in terms of the proportional and level "excess" mortality in winter compared to the rest of the year. We identify 14 causes as having high excess winter mortality.

contamination of the water supply (Jackson et al., 2014; GWPC 2009). There is, at present, no direct evidence on this link between fracking and health (Finkel and Law 2011).[2]

Finally, our empirical strategy is related to that of Myers (2017) who compares households that use heating oil (a petroleum product) or natural gas in Massachusetts to study whether home energy costs are capitalized into home values.

The paper proceeds as follows. We lay out the empirical strategy in the next section, and then describe the data we use to implement it in Section 3. Section 4 presents the results, and Section 5 offers concluding thoughts.

2 Empirical strategy

We estimate the effect of heating prices on mortality. As a proxy for the heating price that an individual faces, we combine information on whether her locality uses natural gas for heating and the national prices of natural gas and electricity. This approach enables us to control for average differences across localities and time.

2.1 Estimating equation

We estimate the following equation via ordinary least squares regression to quantify the impacts of the price of home heating on mortality:

$$log(m_{jt}) = \alpha + \beta_1 ShareGas_j \times log(RelPrice_t) + \gamma_j + \tau_t + \theta Z_j \times log(RelPrice_t) + \delta X_{jt} + \epsilon_{jt} \quad (1)$$

Each observation is a county-bimester (a bimester is a two-month period). The outcome $log(m)$ is the log of age-adjusted mortality in county j in time period t. The key regressor is the interaction of $ShareGas$ — the proportion of people in the area that used natural gas for heating in the base year of 2000 — and $log(RelPrice)$. $RelPrice$ is the ratio of the national price of gas to electricity. When natural gas prices are higher (high $RelPrice$), areas with high $ShareGas$ face relatively higher heating prices. Thus, the hypothesis is that $\beta_1 > 0$, or that a higher heating price increases mortality. County fixed effects γ and bimester-year fixed effects τ absorb the main effects of $ShareGas$ and $log(RelPrice)$. Throughout, we

[2]The health harm from the toxic chemicals used very well might be much larger than the health benefits of lower energy prices per person affected. However, lower energy prices affect a much larger population – hundreds of millions of people across the country – whereas the health harms from chemical contamination are fairly localized. Thus, the net health effect of fracking aggregated for the whole US population is not as clear as one's initial intuition might suggest.

cluster standard errors by state.

We also include several control variables. First, there might be local characteristics Z correlated with $ShareGas$ that have time-varying effects on health with a temporal pattern similar to $RelPrice$. We control for pre-period log median income and the share of the population over age 70 interacted with $log(RelPrice)$. Second, because the study period spans the housing market boom and collapse and the Great Recession, we control for an index of housing prices, the unemployment rate, and the manufacturing share of local employment income. These control variables help safeguard against a spurious correlation, and heating prices could potentially affect firms' labor demand or be incorporated in housing prices (Myers 2017). Finally, the vector X also includes air pollutants that have been linked to mortality and that vary seasonally and across areas.

For certain other outcomes, the available geographic or temporal identifier differs, and the specification differs accordingly. For example, as described in the next section, American Community Survey data only specify the year of the survey, and give the Public Use Micro Area instead of the county. Note that for other outcomes, the unit of observation is a household, that is, household i in geographic unit j and time t.

2.2 Difference-in-difference-in-differences models

For the difference-in-differences estimates given by equation (1), we restrict the data to only winter bimesters (when possible), when energy use is mostly for heating and most of the year's heating is consumed. We can also use the non-winter bimesters as an additional comparison group, estimating a triple difference model:

$$\begin{aligned} log(m)_{jt} &= \alpha + \lambda_1 ShareGas_j \times log(RelPrice_t) \times Winter_t + \lambda_2 ShareGas_j \times log(RelPrice_t) \\ &\quad + \lambda_3 ShareGas_j \times Winter_t + \theta_1 Z_j \times log(RelPrice_t) \times Winter_t \\ &\quad + \theta_2 Z_j \times log(RelPrice_t) + \theta_3 Z_j \times Winter_t + \gamma_j + \tau_t + \delta X_{jt} + \epsilon_{jt} \end{aligned} \tag{2}$$

Here the coefficient of interest is λ_1 which we hypothesize to be positive: The price of heating affects mortality more in winter than in other, warmer months.

Some winters or particular winter months are colder than others, so another approach is to use temperature instead of winter. We do so by constructing for each county-bimester a summary measure of coldness, namely heating degree days, or HDD. HDD sums across a time period (bimester in our case) a daily measure of coldness that is linear in temperature once temperature falls below a threshold, with the convention being 65°F. We can then adapt equation (2) by replacing $Winter$ with HDD:

$$\begin{aligned} log(m)_{jt} &= \alpha + \lambda_1 ShareGas_j \times log(RelPrice_t) \times HDD_{jt} + \lambda_2 ShareGas_j \times log(RelPrice_t) \\ &+\lambda_3 ShareGas_j \times HDD_{jt} + \lambda_4 log(RelPrice_t) \times HDD_{jt} + \lambda_5 HDD_{jt} \\ &+\theta_1 Z_j \times log(RelPrice_t) \times HDD_{jt} + \theta_2 Z_j \times log(RelPrice_t) + \theta_3 Z_j \times HDD_{jt} \\ &+f(\overline{HDD}_j) + \gamma_j + \tau_t + \delta X_{jt} + \epsilon_{jt} \end{aligned} \tag{3}$$

Aside from controlling for the main effect of HDD and its interaction with $RelPrice$, which are not absorbed by bimester fixed effects, equation (3) also controls for the county's average HDD in winter bimesters in parallel to HDD (the $f(\overline{HDD})$ term). Doing so controls for systematic differences (e.g., demographics) between colder regions such as the Midwest and warmer ones such as the South. We show that the results are similar if we omit this control variable, and therefore also use as identifying variation the average difference across places in the severity of their winters.

2.3 Tracing out the income and substitution effects

An auxiliary outcome we examine is the average price of heating experienced by consumers. We calculate the weighted average of the local prices of natural gas and electricity, where weights are the share of households using each heating source. A model analogous to equation (1) but using log average state price as the outcome is like the "first stage" if we were using an instrumental variables approach. We would expect $\beta_1 = 1$ if our regressors were measured without error and if local and national average prices moved entirely in lock-step. The coefficient will be less than 1 if either there is measurement error or there is price variation specific to a locality, which we would expect due to local demand and regulatory factors, and a supply side that is not fully integrated across the US.

We also examine two other outcomes that act as "1.5^{th}" stage equations to get a sense of the importance of the substitution and income effects. First, we examine the (log) quantity of home energy use, combining gas and electricity. When the outcome is log energy use, the coefficient β_1 can be interpreted as a price elasticity. We expect it to be negative, as consumers substitute away from heating when it becomes more expensive. Note that variation in the price of natural gas is mainly measuring variation in a household's heating price while the outcome combines heating plus other energy uses, so the coefficient will be a lower bound on the magnitude of the price elasticity of demand. The data on home energy use do not disaggregate it by purpose (e.g., heating, lighting). Natural gas's home use is mostly for heating (space heating and water heating), with an additional small contribution from stoves. Electricity is the predominant energy source for non-heating home energy needs such as lighting, refrigeration, and air conditioning, throughout the US. Home heating is also the largest home energy use, accounting for 42% of annual home energy consumption in the US, with water heating accounting for an additional 18% (RECS 2009). Other major categories are lighting and appliances (30%), refrigeration (5%), and air conditioning (6%).

To assess the income effect of a higher heating price, we examine expenditures on home energy use, again with the caveat that we cannot distinguish spending on heating from other energy uses (although in winter months, heating accounts for the vast majority of energy use). If the price elasticity of demand is smaller in magnitude than 1, so households are not cutting back one-for-one when the price rises, then we expect higher energy prices to lead to higher energy bills. Of course, we cannot decompose how much of the health effects are due to the changes in the quantity of home heating versus the changes in expenditures on heating bills. As with any price change, the effects we estimate are a bundle of the substitution effect and income effect.

2.4 Background on geographic variation in heating source

Natural gas and electricity are the two most common energy sources used for home heating; they are used by 58% and 30% of households nationwide, respectively. Importantly for our purposes, there is large geographic variation in energy source; in some US communities, almost every household uses natural gas for heating, and in other communities, almost no household uses natural gas. 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, of course, not random, and various factors explain the differences. Natural gas pipelines do not extend to some "corners" of the US, especially sparsely populated ones like Maine. Also, 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 either gas or electricity. Importantly, the geographic differences were determined long before the study period and are highly persistent over time. (The correlation between a county's share using natural gas in 2000 and 2010 is 0.99). Being pre-determined does not rule out that an area's heating source is correlated with other factors related to health, so the analysis controls for other locality characteristics in parallel to heating source.

2.5 Background on temporal variation in energy prices

Figure 2 plots the prices of natural gas and electricity over our 2000 to 2010 study period. The data source is the US Energy Information Administration. (In this figure and throughout the paper, monetary amounts are expressed in 2016 USD.) 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 also rose and then fell over the period.

Because natural gas is one of the fuel sources used in electricity generation, the two prices co-move, just not in lockstep. For example, natural gas prices rose from 2004 to 2005 due in part to supply disruption 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, Medlock III, and Rosthal 2008).

The major reason for the drop in the price of natural gas in the mid-2000s was shale production of natural gas; US shale production of natural gas is also plotted in Figure 2. Advances in hydraulic fracturing of underground shale formations in which natural gas is trapped and in horizontal drilling made shale production economically feasible, and it expanded rapidly beginning around 2006.[3]

2.6 Home heating versus other indoor heating

While we sometimes refer to our results as being due to home heating, the analysis cannot isolate home heating from heating in other buildings such as the workplace. Some policy implications such as whether to promote increased energy supply are similar whether the channel is home heating or other indoor heating. In other cases, for example if one is considering subsidies for consumer heating bills, it would be valuable to be able to isolate

[3]Unlike oil prices, natural gas markets are not fully integrated across the world. Natural gas is easy to ship via pipeline, but not otherwise, and there was a bottleneck in US export capacity in the late 2000s. Thus, US natural gas prices also fell relative to other countries' natural gas prices over this period (Hausman and Kellogg 2015).

heating costs at home, which our research design does not permit. A related limitation is that energy prices affect other energy uses besides home heating. As described above, nationwide, most other major energy uses rely on electricity. An exception is water heating, and thus the results reflect a combination of space and water heating, both of which affect health through similar mechanisms. (Whether a household uses natural gas for water heating is highly correlated with whether it uses it for space heating.)

3 Data

Our empirical analysis focuses on the contiguous US between 2000 and 2010. We exclude Hawaii and Alaska because our data source for temperature does not contain data for them. The rest of this section describes the data sources we use to construct our outcome measures and our key independent variables. Further details are provided in the Data Appendix.

3.1 Mortality

We construct our main outcome, the mortality rate, from restricted-use Vital Statistics microdata. The data consist of records for all deaths in the US, indicating the month and county of residence (and county of death), and the cause of death. The data also include the decedent's age, sex, race, and education level. We conduct our analyses separately for men and women.

We exclude counties with a small population over age 50, specifically those in the bottom tenth percentile of all counties, as they have few (often zero) deaths per month. We aggregate the data to the county-bimester level. Even though the data provide the month of death, we use bimesters to reduce the number of cells with no or few deaths, which arise when we examine specific causes.

Following the literature, we age-adjust the mortality rate using population data from the National Cancer Institute's Surveillance Epidemiology and End Results program. Our main specifications examine the logarithm of the age-adjusted mortality rate, and we also show the results in levels.

Much of our analysis focuses 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, and we zero in on these causes because it is most plausible that they are exacerbated by exposure to cold. To determine these causes, we collapse the data geographically to the entire US, and separately estimate the coefficient of a regression of log age-adjusted mortality on a dummy for winter for each of the National Center for Health Statistics (NCHS) 113 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. This procedure identifies 14 causes that fall within 4 alphabetic (i.e, broad) cause-of-death groups. We exclude one cause, namely deaths from smoke, fire, and flames, because the reason it spikes in winter is quite different in nature to causes with a biological mechanism. The remaining causes generally match the causes highlighted in the literature (e.g., cardiovascular causes and respiratory causes). We add one cause that just misses our criterion for inclusion, cerebrovascular diseases (strokes), as this cause has been linked in the literature to winter and cold weather (The Eurowinter Group 1997; Sheth et al. 1999). Appendix Figure 3 shows the seasonality for our EWM causes and for other causes; one noteworthy pattern is that even the non-EWM causes exhibit noticeable seasonality for women, so EWM causes capture all of the relevant effects a bit better for men than women. Appendix Table A1 lists the causes and their degree of EWM.

3.2 Independent variables

To construct $ShareGas$, we use 2000 Decennial Census data. The Census longform asks the energy source for home heating, as does the American Community Survey (ACS), which has been fielded annually since 2005. We use the 2000 Decennial Census of Population and Housing Summary Files that report aggregate data for each census tract or county or state. When our geographic unit is the Public Use Micro Area, or PUMA, we construct each PUMA's value of $ShareGas$ from public use microdata.

$RelPrice$, the ratio of the national price of gas to electricity, is constructed using the monthly national prices of natural gas and electricity for residential customers available from the US Energy Information Administration (EIA). The EIA monthly data is based on a survey of a sample of utility companies; to minimize measurement error, we average prices over the past six months.[4] Price is defined in the data as the total revenue divided by total sales volume of natural gas or electricity supplied to residential consumers, and the data are

[4] The EIA also does a survey of the universe of firms, which yields an annual-level prices dataset. We obtain similar results if we use this dataset to construct $RelPrice$.

available at the national and state level.

The analysis also uses temperature data. We start with the PRISM dataset of daily average temperature for gridpoints across the contiguous US spaced 4 kilometers apart (PRISM Climate Group 2004). We calculate the temperature for each census block and then use population weighting to construct the average for each county (or other geographic unit). Our observations are at the bimester level, so we use the daily data to construct heating degree days (HDD) for the bimester. HDD is a commonly used measure of coldness or need for heating that is based on the idea that demand for heating is linear in temperature once temperature falls below some threshold. That is, $HDD_{jt} = \sum_{x=1}^{T} \max\{threshold - tmean_{jx}, 0\}$ where $threshold$ is a temperature threshold, $tmean$ is the mean temperature of the area j on day x of bimester t, and T is the number of days in bimester t. We use the conventional threshold of 65°F. For ease of interpretation of the regression coefficients, we normalize the measure so that an increase of 1 unit in HDD is the difference between a temperature of 65°F or above each day in the month versus 32°F each day.

We also use several control variables. First, we use air pollution from the Air Quality System (AQS) of the US Environmental Protection Agency (EPA). Air pollution is an important control variable because it is correlated with weather conditions and affects mortality (Ye et al. 2012). We aggregate the daily monitoring-station-level air quality indices (AQI) for particulate matter (2.5 micron and 10 micron), carbon monoxide, nitrogen dioxide, ozone, and sulfur dioxide to the county-bimester level. In our main specifications we focus on particulate matter and nitrogen dioxide, as these are the pollutants correlated with mortality.

We also control for other area characteristics in parallel to $ShareGas$. Specifically, we use 2000 Census data and calculate the median household income and the proportion of the population that is age 70 and older. In addition, because a major housing market run-up and collapse occurred during the study period, we control for a Housing Price Index (HPI), available at the state-quarter level through the Federal Housing Finance Agency. Similarly, the 2009 recession had different impacts across counties, and, in addition, the price of heating might have an impact on the local economy. We therefore control for the unemployment rate (available at the county-month level from the Bureau of Labor Statistics) and the manufacturing sector share of total employee compensation (available at the state-quarter level form the Bureau of Economic Analysis).

3.3 Other dependent variables

We examine intermediate outcomes to shed light on why heating prices affect mortality. As discussed in the previous section, our "first stage" outcome is the local price of home energy, which we compute as the consumption-weighted average of residential natural gas and electricity prices. To assess how price affects heating usage, we examine the impact on total residential energy use, which we compute as the sum of natural gas and electricity. Price and usage data are aggregate state-month-level data from EIA.[5]

We use two data sources to measure household spending on home energy. First, we combine 2000 Census microdata (IPUMS version) and ACS data for 2005 to 2010. This analysis is conducted at the PUMA level rather than county, as the PUMA is the finest geographic identifier provided in the data. While this data source has a gap from 2001 to 2004 (because the ACS only began in 2005), it is a very large data set with over one million observations per year. For simplicity and computational ease, we collapse the data to the PUMA-year level for the analysis. Second, we use the Consumer Expenditure Survey (CEX). The CEX is available on a rolling-quarter basis but only provides the respondent household's state of residence and is not available for all states. The CEX allows us to confirm the income effect, and that it is concentrated in winter, but we have limited statistical power to examine other expenditure patterns.

4 Results

In this section, we first present our results on the "first stage" and "1.5$^{\text{th}}$" stage outcomes of home energy prices, quantity of energy consumed, and energy bills. We then turn to our main results: the impacts on mortality. Finally, to shed further light on any channels that occur through an income effect, we examine other spending by the household.

4.1 Effect of heating prices on usage and bills

In our analysis, we use $ShareGas \times log(RelPrice)$ as an exogenous source of variation in the home heating prices faced by families. Our data sets with our main outcomes do not collect data on households' energy prices – only their energy expenditures. However, we

[5]Natural gas prices and quantities are provided on a volumetric basis in the data. When we construct average energy prices or total energy consumption, to make the prices comparable, we use additional data from the EIA on the BTU content of natural gas supplied to residential customers, available at the state-firm-year level.

are able to use aggregate administrative data on residential energy prices to verify that our regressor is a good proxy for household prices.

In the analysis, the outcome is the weighted average price of residential natural gas and electricity prices, and each observation is a state-month. As shown in Table 1, columns 1 to 3, home energy prices are strongly correlated with $ShareGas \times log(RelPrice)$. The columns sequentially add control variables to the specification. In column 1, we include state and month-year fixed effects. In column 2, we add the housing price index and the interactions of $log(RelPrice)$ with median income and the share of people over age 70. In column 3, we include the rest of the control variables discussed above: air pollution, the unemployment rate, and the manufacturing share of the economy.

The fact that the coefficient is less than 1 is due to several factors. First and foremost, the outcome is average *energy* prices, while the regressor is intended to proxy for average *heating* prices. Second, the outcome is average prices weighted by usage, so it also incorporates any responses of usage to prices. Third, we construct $ShareGas$ weighting each household equally, whereas the usage-weighted measure from EIA implicitly weights bigger users (e.g., larger dwellings) more. Fourth, there might be some measurement error in $ShareGas$. Note that when we examine effects on mortality, the relevant scale factor is 1, not the smaller coefficient estimated here; a change in $ShareGas \times log(RelPrice)$ can still be interpreted as a proportional change in the heating price faced by a household.

We next quantify how households' energy use responds to higher prices and the impact on their 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.) We start by using the EIA data to examine the impact on usage, shown in Table 1, columns 4 to 6. As expected, higher prices lead to less consumption. Both the outcome and regressor are given in logs, so the coefficient has the units (or, more precisely, lack of units) of an elasticity. The coefficient of about -0.4 implies that households are cutting back usage quite a bit, but not one-for-one with prices. In this case, to quantify the elasticity, it is best to scale the coefficient by the corresponding price-change coefficient from columns 1 to 3. We report this implied elasticity (-0.4 to -0.5) at the bottom of the table.

Given that the elasticity appears to be smaller in magnitude than -1, households are spending more money on energy expenses when the heating price increases. We can verify that this is the case using household level data from two data sources, Census/ACS and CEX. Columns 1 to 5 of Table 2 show the results using the Census/ACS. We see that the heating price shock variable is associated with a 0.34 log point increase in energy expenses (columns 1 to 3). We do not know the month of the ACS survey, so this coefficient underestimates the impact during winter months if the result is mainly driven by changes in winter expenses. (Note that ACS also does not release the distribution of when surveys are completed. Census data are all collected at one date but the question on energy bills asks about annual spending.) Column 4 examines the outcome in levels, and a 10% increase in heating prices is associated with a 6.8 USD increase in the home energy bill, averaged over the year. Another useful benchmark is the 25% (0.29 log point) decrease in the relative price of natural gas between 2006 and 2010. This price decline corresponds to a $17 decline in the monthly energy bill, averaged over the year or 3*$17 = $51 per month if driven by November to January. The mean monthly energy bill is $232; hence, the price decline would have induced a decrease of around 7% in the annual energy bill. Column 5 shows the effect as a percent of income. As reported at the bottom of the table, households spend on average 7 percent of income on energy bills. A 25% decline in the heating price leads to a 0.8 percentage point decrease in the proportion of income spent on energy bills.

We find similar results for home utilities expenditures in winter months using CEX data (columns 6 and 7). The elasticity of total utility spending with respect to the heating price is 0.44. The coefficient of $136 in column 7 implies that the 25% price decline between 2006 and 2010 corresponds to a $34 per month decline in winter utilities expenditure, which is in the same range as we found using ACS data.

To summarize, we find that home heating use is quite price elastic, with households cutting back usage in response to our measure of the heating price shock. At the same time, they do not cut back fully and thus also experience a meaningful increase in energy bills when the price of heating increases.

4.2 Effect of heating prices on mortality

We now turn to examining the impact of heating prices on mortality. We report separate results for males and females. We use the log of the age-adjusted mortality rate as the main outcome, following Stevens et al. (2015), and also show robustness to using the age-adjusted rate.

Table 3 presents the results on all-cause mortality. Column 1 includes as regressors only the basic area and time fixed effects in addition to $ShareGas \times log(RelPrice)$. Column 2 controls for median income and the proportion of the population age 70 and above in parallel to $ShareGas$ and the housing price index. Column 3 adds in air pollution controls, the unemployment rate, and the manufacturing share of the economy. The top panel shows that for males, the estimated elasticity of mortality with respect to the heating price is 0.07 and 0.10. The bottom panel shows the effect for women. We find smaller point estimates for women (throughout the analysis), but based on fully interacted regressions (not reported), we generally cannot reject equality between the male and female estimates. The last three columns of the table reproduce the results with the age-adjusted mortality rate in levels, and we see consistent patterns.

For the remainder of the analysis we focus on high-EWM causes of death, that is, causes selected as having a pronounced seasonal pattern with peak mortality in winter. While, as a bottom line estimate, we might care more about total mortality, focusing on high-EWM causes is a more honed test of the hypothesis that heating prices affect deaths due to exposure to cold and gives us more statistical power, for example to test for heterogeneous effects. Table 4 is structured in the same way as Table 3, but the outcome is EWM mortality. For males, we find a strong positive effect of the price of heating on mortality. For females, the effects are smaller in magnitude and, in our preferred specification (column 3), insignificant. The fact that the female all-cause effect looks fairly comparable to the male effect, but the EWM effect does not reflects the fact that even the causes not selected as having high-EWM are more peaked in winter for women than men (as seen in Appendix Figure 3).

We next turn to examining effects by cause of death. We start by aggregating the 14 narrow causes into the alphabetic ICD categories they fall within: non-viral non-respiratory infections; neurological; circulatory; and respiratory. Table 5 presents the results, and we see that the overall effect for men is mainly driven by circulatory and respiratory causes. As the means presented at the bottom of the table show, circulatory and respiratory causes are much larger contributors to mortality overall; it is thus difficult to say whether the proportional effect is smaller for the other two categories or we have less power to detect effects. The last four columns show the results for women. Here we only find a marginally significant effect for respiratory causes and no significant effect of heating prices on other causes of mortality. We provide further detail on the breakdown by narrow cause by examining results separately for each of the 14 EWM causes, reported in Appendix Table A2.

We next turn to our difference-in-difference-in-differences estimates that bring in data from non-winter months. We use as the third difference either $Winter$ or HDD (a summary measure of average temperature). Given that the overall effects are concentrated among males, we show the results for males, and present the (mostly null) results for females in the appendix. Table 6, columns 1 to 3, show that the effect for males is stronger in winter than the rest of the year.[6] The fact that the coefficients are very similar to those shown in Table 5 means that when we run our specifications using only non-winter months as a placebo test, reassuringly, the price of heating has no effect on mortality.

We similarly find that the price of heating increases mortality more in colder bimesters, as measured by HDD, as reported in Table 6, columns 4 to 6. HDD is scaled so that a unit change is the difference between every day in the month being 65°F or above and being 32°F. The specification controls for the average winter HDD for the county in parallel to HDD, to remove systematic differences in mortality between places that happen to have cold versus warm winters. Doing so might be viewed as overcontrolling, so we also report the results without controlling for average HDD in Appendix Table A4. The results are similar but somewhat weaker, which is also consistent with results in the literature that due to adaptation (e.g., better insulated homes in colder places), atypical cold for an area is what affects mortality more (The Eurowinter Group 1997). Indeed, if we correlate HDD and mortality (ignoring heating prices), HDD is more predictive of mortality conditional on average HDD than unconditionally.

We also report results in Appendix Table A4 for an alternative measure, $HDD32$, which uses 32°F as the threshold. While 65°F is the convention for heating demand, one might think that it is the very cold days that affect mortality; a temperature of 25°F might increase mortality, but maybe a 55°F day does not, even if people use heating. We find similar results for $HDD32$. The interquartile range of $HDD32$ is one third as large as the interquartile range for HDD, so to compare coefficients, one would want to divide the $HDD32$ coefficients by 3. That calculation suggests slightly bigger effects using the $HDD32$ measure. We use HDD as our preferred measure both because it is the convention in the literature, and because HDD is more predictive of mortality than $HDD32$ (in a regression that ignores heating prices and aims to simply estimate the temperature-mortality relationship).

[6] The corresponding results for females are presented in Appendix Table A3.

5 Conclusion

In this paper, we estimate the causal effect of heating prices on mortality in the US. We find that winter mortality is lower when the price of heating is lower: the elasticity of mortality with respect to the price of heating is 0.10 for males and 0.08 for females. Our difference-in-difference-in-differences models makes comparisons across counties, years, and either seasons of the year or months when it is colder versus less cold. Thus, any time-varying factors specific to areas using natural gas for heating that could generate bias would also have to be seasonal in nature or correlated with contemporaneous temperature.

To put the elasticities that we estimate in context, the price of natural gas relative to electricity fell by 25% between 2006 to 2010, largely due to the boom in shale gas production. This relative price decline led to a 2% to 3% decrease in the mortality rate for households using natural gas for heating. Given that 58% of American households use natural gas for heating, overall, lower natural gas prices reduced winter mortality rate by 1.5% over the time period.

Given the large share of income spent on energy bills—7% on average but considerably higher for poorer households—we should not be surprised that households are reacting to changes in the heating price by adjusting their expenditure patterns. We find that households respond to higher heating prices by consuming less energy in winter months (i.e. there is a substitution effect), but with an elasticity smaller in magnitude than -1. We verify that households are spending more money on energy expenses when the heating price increases (i.e. there is also an income effect). Either one of the two effects could be the channel through which the mortality rate increases, and we leave assessing the importance of each mechanism for future work.

In ongoing analyses, we are examining the effect of heating prices on health outcomes other than mortality, such as physical morbidities and mental health, using the Health and Retirement Study and other data sources.

References

Analitis, A., K Katsouyanni, A Biggeri, M Baccini, B Forsberg, L Bisanti, U Kirchmayer, F Ballester, E Cadum, P. Goodman, et al. (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.

Bartik, A. W., J. Currie, M. Greenstone, and C. R. Knittel (2017). “The Local Economic and Welfare Consequences of Hydraulic Fracturing”. (23060).

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.

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.

DeLeire, T., P. Eliason, and C. Timmins (2014). “Measuring the Employment Impacts of Shale Gas Development”. Unpublished Manuscript, McCourt School of Public Policy, Georgetown University.

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

Feyrer, J., E. T. Mansur, and B. Sacerdote (2017). “Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution”. *The American Economic Review* 107 (4), pp. 1313–1334.

Finkel, M. L. and A. Law (2011). “The Rush to Drill for Natural Gas: A Public Health Cautionary Tale”. *American Journal of Public Health* 101 (5), pp. 784–785.

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, et al. (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.

Hartley, P. R., K. B. Medlock III, 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*.

Myers, E. (2017). “Are Home Buyers Myopic? Evidence From Capitalization of Energy Costs”.

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

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

Sheth, T., C. Nair, J. Muller, and S. Yusuf (1999). "Increased Winter Mortality From Acute Myocardial Infarction and Stroke: The Effect of Age". *Journal of the American College of Cardiology* 33 (7), pp. 1916–1919.

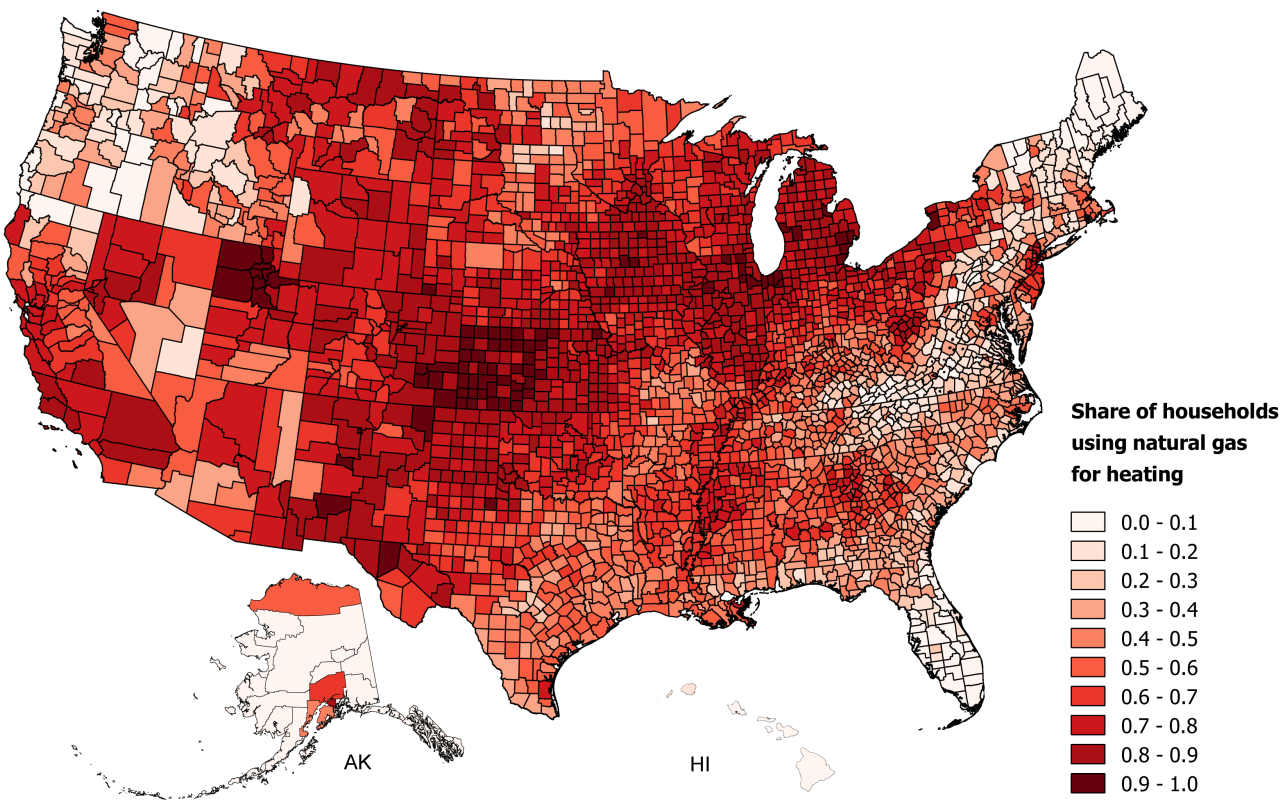
Stevens, A. H., D. L. Miller, M. E. Page, and M. Filipski (2015). "The Best of Times, the Worst of Times: Understanding Pro-Cyclical Mortality". *American Economic Journal: Economic Policy* 7 (4), pp. 279–311.

The 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.

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.

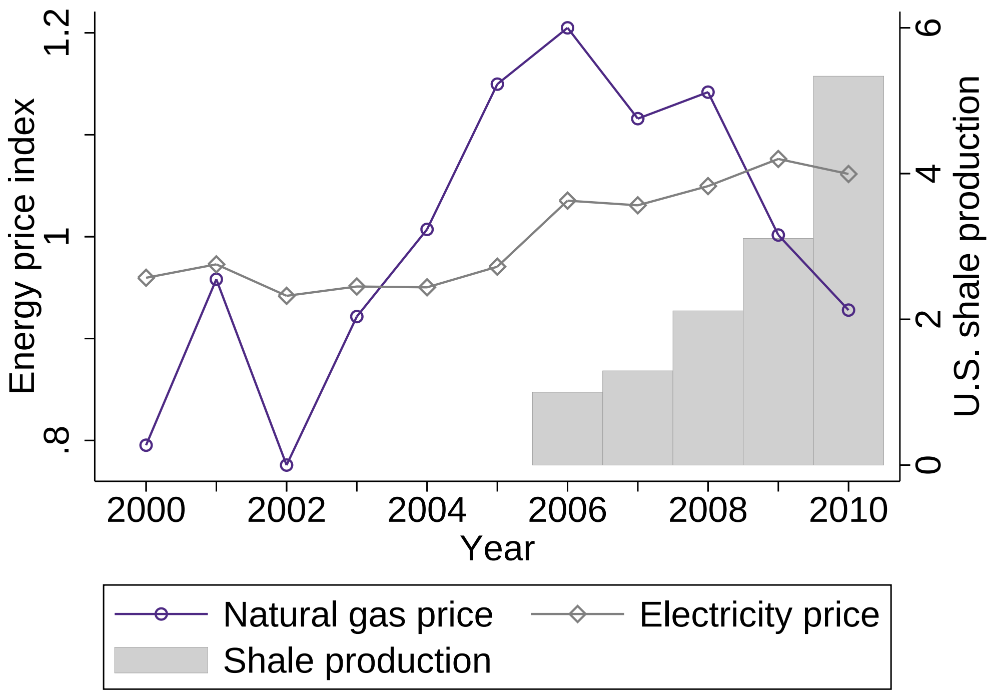
Ye, X., R. Wolff, W. Yu, P. Vaneckova, X. Pan, and S. Tong (2012). "Ambient Temperature and Morbidity: A Review of Epidemiological Evidence". *Environmental Health Perspectives* 120 (1), p. 19.

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



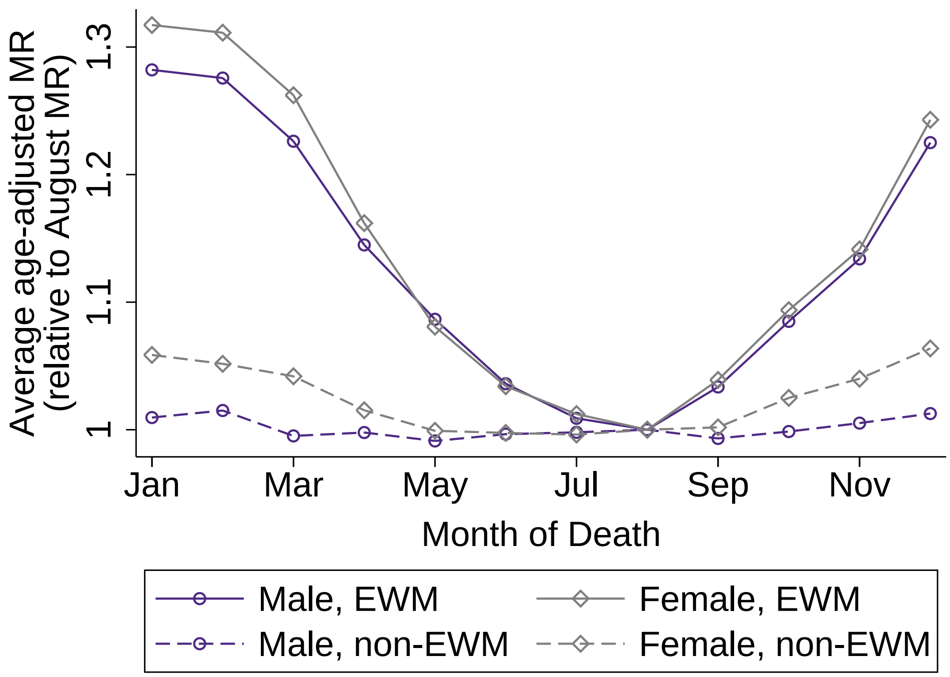
Notes: Data are from the 2000 US Census.

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 in the background (right axis). Data are from EIA.

Figure 3: Seasonality in mortality



Notes: Average age-adjusted mortality rates across US counties (excluding Alaska) between 2000 and 2010, broken down by sex and high-EWM and other causes. High-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. We normalize each series by its value in August (the month with the lowest all-cause mortality rate). Data are Vital Statistics data from NCHS.

Table 1: Effect of the price of heating 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)	(5)	(6)
ShareGas $\times$ Log(RelPrice)	0.75***	0.71***	0.78***	-0.37*	-0.29	-0.36*
	[0.15]	[0.16]	[0.16]	[0.18]	[0.19]	[0.19]
Observations	2,156	2,156	2,156	2,156	2,156	2,156
Mean price/quantity	21.1	21.1	21.1	22.6	22.6	22.6
Basic fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Basic controls	No	Yes	Yes	No	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes
Implied elasticity				-0.49	-0.41	-0.46

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of state-year-months on the contiguous U.S. for winter months (November-February) 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. *ShareGas* is the proportion of occupied housing units in each state in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past six months, to the corresponding residential price of electricity. Monetary variables are in constant 2016 dollars. Basic fixed effects are state and year-month fixed effects. Basic controls are the interactions of *Log(RelPrice)* with the log of the median state household income in 1999 and the share of people aged 70 and above in 2000, and the state housing price index. Additional controls are all other variables used in our main specification: the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, and the state's manufacturing sector share of total employee compensation. Implied elasticity is the ratio of the coefficient to the corresponding coefficient from the first three columns (the "first stage").

Table 2: Effect of the price of heating on home energy bills and expenditure

	Dependent variable:						
	Log of total energy bill	Log of total energy bill	Log of total energy bill	Total energy bill	Energy bill as % of HH income	IHS of utilities expenditure	Utilities expenditure
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ShareGas × Log(RelPrice)	0.35***	0.34***	0.34***	68.3***	3.02***	0.44*	135.8***
	[0.084]	[0.094]	[0.089]	[23.0]	[0.69]	[0.24]	[42.1]
Observations	14,385	14,385	14,385	14,385	14,385	43,772	43,772
Mean bill/percentage/expenditure	231.6	231.6	231.6	231.6	7.20	314.5	314.5
Basic fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Basic controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	No	No	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Columns 1 to 5: The sample is composed of PUMA-years on the contiguous U.S. included in the 2000 Census and the ACS PUMS data between 2005 and 2010. Total energy bill is the mean monthly bill from electricity, gas and other fuels in the PUMA. Columns 6 and 7: The sample is composed of households on the contiguous U.S. interviewed in February or March between 2000 and 2010 in the CEX Interview Survey. Utilities expenditures are average monthly expenditures on utilities, fuels and public services for the three months prior to interview. We drop the top 5% of expenditure values. IHS refers to the inverse hyperbolic sine transformation. *ShareGas* is the proportion of occupied housing units in each PUMA/state in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of annual average residential price of natural gas to annual average residential price of electricity in the US. Basic fixed effects are PUMA and year fixed effects for columns 1 to 5, and state and year-month (interview month) fixed effects for columns 6 and 7. Basic controls are the interaction of *Log(RelPrice)* with the log of the median PUMA/state household income in 1999 and the share of people aged 70 and above in 2000, and the state housing price index. Additional controls are all other variables used in our main specification: the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, and the state's manufacturing sector share of total employee compensation. Monetary variables are in constant 2016 dollars.

Table 3: Effect of the price of heating on all-cause mortality

	Dependent variable: Log of all-causes MR			Dependent variable: All-causes MR		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Deaths in male population.			
ShareGas × Log(RelPrice)	0.071*	0.093**	0.10**	112.9**	129.2***	128.6***
	[0.039]	[0.040]	[0.039]	[45.0]	[45.1]	[45.8]
Observations	61,319	61,319	61,319	61,336	61,336	61,336
Mean MR	1119.7	1119.7	1119.7	1119.4	1119.4	1119.4
			Panel B: Deaths in female population.			
ShareGas × Log(RelPrice)	0.060	0.084*	0.081*	70.8**	80.1**	72.5**
	[0.043]	[0.045]	[0.043]	[32.0]	[35.4]	[33.5]
Observations	61,310	61,310	61,310	61,336	61,336	61,336
Mean MR	786.9	786.9	786.9	786.6	786.6	786.6
Basic fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Basic controls	No	Yes	Yes	No	Yes	Yes
Pollution controls	No	No	Yes	No	No	Yes

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. for winter bimesters (November-February) between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. Basic fixed effects are county and year-bimester fixed effects. Basic controls are the interaction of *Log(RelPrice)* with the log of the median county household income in 1999, the interaction of *Log(RelPrice)* with the share of people aged 70 and above in 2000, and the state housing price index. Additional controls are the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, and the state's manufacturing sector share of total employee compensation.

Table 4: Effect of the price of heating on mortality from high-EWM causes of death

	Dependent variable: Log of EWM causes MR			Dependent variable: EWM causes MR		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Deaths in male population.						
ShareGas × Log(RelPrice)	0.16^{***}	0.17^{***}	0.17^{***}	127.3^{***}	134.8^{***}	124.5^{***}
	[0.048]	[0.052]	[0.052]	[36.0]	[38.8]	[38.3]
Observations	61,185	61,185	61,185	61,336	61,336	61,336
Mean MR	710.0	710.0	710.0	708.2	708.2	708.2
Panel B: Deaths in female population.						
ShareGas × Log(RelPrice)	0.059	0.091^{*}	0.092^{*}	51.6^{**}	60.3^{**}	54.1^{**}
	[0.052]	[0.049]	[0.047]	[22.3]	[23.7]	[22.6]
Observations	61,162	61,162	61,162	61,336	61,336	61,336
Mean MR	504.5	504.5	504.5	503.1	503.1	503.1
Basic fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Basic controls	No	Yes	Yes	No	Yes	Yes
Pollution controls	No	No	Yes	No	No	Yes

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. for winter bimesters (November-February) between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. High-EWM causes are those that exhibit high mortality in winter compared to the rest of the year. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. Basic fixed effects are county and year-bimester fixed effects. Basic controls are the interaction of *Log(RelPrice)* with the log of the median county household income in 1999, the interaction of *Log(RelPrice)* with the share of people aged 70 and above in 2000, and the state housing price index. Additional controls are the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, and the state's manufacturing sector share of total employee compensation.

Table 5: Effect of the price of heating on mortality, by ICD-10 category

	Dep. var.: Log of male EWM-causes MR in ICD-10 group				Dep. var.: Log of female EWM-causes MR in ICD-10 group			
	Group A: Non-viral, non-respiratory infections	Group G: Neurological diseases	Group I: Circulatory system diseases	Group J: Respiratory system diseases	Group A: Non-viral, non-respiratory infections	Group G: Neurological diseases	Group I: Circulatory system diseases	Group J: Respiratory system diseases
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ShareGas × Log(RelPrice)	0.12	0.069	0.21***	0.21***	0.070	0.13	0.064	0.16***
	[0.089]	[0.084]	[0.060]	[0.062]	[0.084]	[0.085]	[0.051]	[0.055]
Observations	42,679	40,229	60,716	59,573	44,369	47,324	60,671	59,293
Mean MR	85.52	83.56	473.7	327.9	66.78	68.65	324.8	214.0

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. for winter bimesters (November-February) between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. All columns include as covariates basic fixed effects, basic controls, and additional controls from Tables 3 and 4, i.e., county and year-bimester fixed effects; *Log(RelPrice)* interacted with the log of the median county household income in 1999, *Log(RelPrice)* interacted with the share of people aged 70 and above in 2000, and the state housing price index; and the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, and the state's manufacturing sector share of total employee compensation.

Table 6: DDD estimates using winter and temperature (male subsample)

	Dependent variable: Log of male EWM-causes MR in ICD-10 group					
	All groups	Group I: Circulatory system diseases	Group J: Respiratory system diseases	All groups	Group I: Circulatory system diseases	Group J: Respiratory system diseases
	(1)	(2)	(3)	(4)	(5)	(6)
ShareGas × Log(RelPrice) × Winter	0.17***	0.21***	0.15**			
	[0.055]	[0.065]	[0.060]			
ShareGas × Log(RelPrice) × HDD				0.19**	0.31***	0.22***
				[0.079]	[0.093]	[0.073]
Observations	183,358	181,742	177,629	183,358	181,742	177,629
Mean MR	651.4	437.5	297.2	651.4	437.5	297.2

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. *Winter* is a binary variable that takes on a value of one in winter bimesters (November to February). *HDD* is the number of heating degree-days in the county for the bimester, based on a threshold of 65°F, in units of °F-days divided by 2000, and scaled to a 60-day bimester. All columns include as covariates county and year-bimester fixed effects. The first three columns include all possible two-way interactions between *ShareGas*, *Log(RelPrice)*, and *Winter*, unless collinear with year-bimester fixed effects. The last three columns analogously include all two-way interactions between *ShareGas*, *Log(RelPrice)*, and *HDD*, and additionally include *HDD*. The last three columns also include the interaction of the average county HDD in winter bimesters with *Log(RelPrice)*, and the three-way interactions of the average county HDD in winter bimesters, *Log(RelPrice)*, and each of *ShareGas*, the log of the median county household income in 1999 and the share of people aged 70 and above in 2000. All columns also include the state housing price index, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, the state's manufacturing sector share of total employee compensation, and the two- and three-way interactions among *Log(RelPrice)*, either *Winter* or *HDD*, and each of the log of the median county household income in 1999 and the share of people aged 70 and above in 2000.

Appendix Table A1: 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	0.15	0.15
Parkinson's disease (G20-G21)	0.53	0.08	0.15
Alzheimer's disease (G30)	1.92	0.33	0.17
Acute myocardial infarction (I21-I22)	4.34	0.66	0.14
Atherosclerotic cardiovascular disease (I25.0)	1.77	0.25	0.13
All other forms of chronic ischemic heart disease (I20,I25.1-I25.9)	6.32	0.83	0.13
Heart failure (I50)	1.61	0.21	0.13
Cerebrovascular diseases (I60-I69)	4.12	0.53	0.12
Atherosclerosis (I70)	0.30	0.05	0.15
Influenza (J09-J11)	0.04	0.06	2.08
Pneumonia (J12-J18)	1.63	0.61	0.34
Emphysema (J43)	0.38	0.09	0.21
Other chronic lower respiratory diseases (J44,J47)	3.11	0.62	0.19
Pneumonitis due to solids and liquids (J69)	0.47	0.09	0.17
Other diseases of respiratory system (J00-J06,J30- J39,J67,J70-J98)	0.77	0.10	0.13
Accidental exposure to smoke, fire and flames (X00-X09)	0.09	0.06	0.58

Notes: Mortality rates are expressed per 100,000 population. The 75th percentile of level and log coefficient are 0.02 and 0.13, respectively. We add *Cerebrovascular diseases* and remove *Accidental exposure to smoke, fire and flames* from the data-driven definition of excess winter mortality causes.

Appendix Table A2: Effect of the price of heating on mortality, by specific cause of death

	Dependent variable: Log of specified disease MR among:				
	Males	Females		Males	Females
	(1)	(2)		(1)	(2)
Septicemia	0.12	0.070	Atherosclerosis	0.22**	0.11
	[0.089]	[0.084]		[0.11]	[0.15]
	(42,679)	(44,369)		(27,568)	(30,012)
	{85.5}	{66.8}		{55.2}	{39.9}
Parkinson's disease	0.060	0.12	Influenza	-0.31	-0.73
	[0.094]	[0.11]		[0.42]	[0.49]
	(25,459)	(22,664)		(3,452)	(4,189)
	{46.2}	{24.5}		{30.3}	{22.8}
Alzheimer's disease	0.12	0.15	Pneumonia	0.21***	0.35***
	[0.084]	[0.10]		[0.075]	[0.082]
	(33,971)	(44,989)		(50,614)	(50,998)
	{67.0}	{61.1}		{132.2}	{87.4}
Acute myocardial infarction	0.21*	0.20**	Emphysema	0.40**	0.11
	[0.11]	[0.078]		[0.18]	[0.15]
	(52,322)	(49,371)		(21,880)	(18,158)
	{140.0}	{82.5}		{38.2}	{24.3}
Chronic ischemic heart disease	0.24***	0.090	Oth. chronic lower resp. diseases	0.31***	0.17**
	[0.067]	[0.072]		[0.084]	[0.077]
	(57,223)	(55,591)		(53,276)	(50,716)
	{241.3}	{138.9}		{149.0}	{91.4}
Heart failure	0.29***	0.12	Pneumonitis (solids and liquids)	0.13	0.13
	[0.090]	[0.077]		[0.12]	[0.16]
	(53,906)	(56,135)		(31,646)	(30,008)
	{160.8}	{121.6}		{60.2}	{35.2}
Cerebrovascular diseases	0.22***	0.20**	Other respiratory diseases	0.051	0.24***
	[0.077]	[0.083]		[0.076]	[0.083]
	(50,749)	(54,107)		(50,243)	(49,919)
	{127.3}	{106.2}		{128.6}	{92.7}

Notes: Each cell shows the result from a separate regression, and reports the coefficient on *ShareGas* × *Log(RelPrice)*, the corresponding standard error clustered by state in brackets, the number of observations in parentheses, and the mean mortality rate of the specified cause in braces. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. for winter bimesters (November-February) between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. All columns include as covariates basic fixed effects, basic controls, and additional controls from Tables 3 and 4.

Appendix Table A3: DDD estimates using winter months and temperature (female subsample)

	Dependent variable: Log of female EWM-causes MR in ICD-10 group					
	All groups	Group I: Circulatory system diseases	Group J: Respiratory system diseases	All groups	Group I: Circulatory system diseases	Group J: Respiratory system diseases
	(1)	(2)	(3)	(4)	(5)	(6)
ShareGas × Log(RelPrice) × Winter	0.061	-0.039	0.084			
	[0.049]	[0.056]	[0.057]			
ShareGas × Log(RelPrice) × HDD				0.023	-0.093	0.046
				[0.080]	[0.087]	[0.082]
Observations	183,345	181,685	176,063	183,345	181,685	176,063
Mean MR	459.6	299.3	191.9	459.6	299.3	191.9

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. *Winter* is a binary variable that takes on a value of one in winter bimesters (November to February). *HDD* is the number of heating degree-days in the county for the bimester, based on a threshold of 65°F, in units of °F-days divided by 2000, and scaled to a 60-day bimester. All columns include as covariates county and year-bimester fixed effects. The first three columns include all possible two-way interactions between *ShareGas*, *Log(RelPrice)*, and *Winter*, unless collinear with year-bimester fixed effects. The last three columns analogously include all two-way interactions between *ShareGas*, *Log(RelPrice)*, and *HDD*, and additionally include *HDD*. The last three columns also include the interaction of the average county HDD in winter bimesters with *Log(RelPrice)*, and the three-way interactions of the average county HDD in winter bimesters, *Log(RelPrice)*, and each of *ShareGas*, the log of the median county household income in 1999 and the share of people aged 70 and above in 2000. All columns also include the state housing price index, the AQIs for $PM_{2.5}$, PM_{10}, and NO_2, the unemployment rate, the state's manufacturing sector share of total employee compensation, and the two- and three-way interactions among *Log(RelPrice)*, either *Winter* or *HDD*, and each of the log of the median county household income in 1999 and the share of people aged 70 and above in 2000.

Appendix Table A4: DDD estimates with heating degree days using alternative specifications

	Dependent variable: Log of male EWM-causes MR in ICD-10 group			Dependent variable: Log of female EWM-causes MR in ICD-10 group		
	All groups	Group I: Circulatory system diseases	Group J: Respiratory system diseases	All groups	Group I: Circulatory system diseases	Group J: Respiratory system diseases
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Without controlling in parallel for average winter HDD.						
ShareGas × Log(RelPrice) × HDD	0.11	0.19**	0.15**	0.025	-0.053	0.062
	[0.067]	[0.086]	[0.061]	[0.069]	[0.070]	[0.084]
Panel B: HDD defined based on threshold of 32°F.						
ShareGas × Log(RelPrice) × HDD	0.83**	1.27***	0.19	-0.10	-0.42	-0.46
	[0.36]	[0.35]	[0.30]	[0.35]	[0.44]	[0.37]
Observations	183,358	181,742	177,629	183,345	181,685	176,063
Mean MR	651.4	437.5	297.2	459.6	299.3	191.9

Notes: Standard errors clustered by state in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample is composed of county-year-bimesters on the contiguous U.S. between 2000 and 2010, excluding counties with population aged 50 and over in 2000 less than the 10th percentile among all counties. Mortality rates are age-adjusted mortality rates, computed using NCHS mortality data; see data appendix for further details. *ShareGas* is the proportion of occupied housing units in each county in 2000 that indicate that gas is their main heating source. *Log(RelPrice)* is the log of the ratio of the monthly residential price of natural gas in the US, averaged over the past three bimesters, to the corresponding residential price of electricity. *HDD* is the number of heating degree-days in the county for the bimester, based on thresholds of 65°F and 32°F in panels A and B respectively, in units of °F-days divided by 2000, and scaled to a 60-day bimester. All columns include as covariates county and year-bimester fixed effects; *HDD*; all possible two-way interactions between *ShareGas*, *Log(RelPrice)*, and *HDD*, unless collinear with year-bimester fixed effects; the state housing price index; the AQIs for $PM_{2.5}$, PM_{10}, and NO_2; the unemployment rate, the state's manufacturing sector share of total employee compensation, the two- and three-way interactions among *Log(RelPrice)*, *HDD*, and each of the log of the median county household income in 1999 and the share of people aged 70 and above in 2000. Panel B additionally include the interaction of the average county HDD (defined based on a threshold of 32°F) in winter bimesters with *Log(RelPrice)*, and the three-way interactions of the average county HDD in winter bimesters, *Log(RelPrice)*, and each of *ShareGas*, the log of the median county household income in 1999 and the share of people aged 70 and above in 2000.