The Adaptation of the US Residential Sector to Global Warming

François Cohen, Grantham Research Institute, London School of Economics

Matthieu Glachant*, MINES ParisTech

Magnus Söderberg, University of Gothenburgh

April, 2016

Abstract

Good quality homes will be necessary to protect households from unusually elevated temperatures caused by climate change. Using household-level data from the American Housing Survey, this paper assesses the economic cost of adaptation to higher temperatures in the residential sector through home improvements and changes in energy consumption. Our best estimate of the present discounted value of the cost for adapting to temperature increases under the "business-as-usual" A2 climate change scenario is $3,950 per housing unit, but the null hypothesis of a zero cost cannot be rejected at conventional significance levels. To put this in context, this is equivalent to an annual expenditure of 210 $, which is approximatively 0.2% of the sample average annual household income or about 1.9% of the average purchase price of the housing units included in the sample (around 211,000 $). So this appears to be a small loss.

JEL Codes: D12, Q47, Q54, R22.

Keywords: Climate change, adaptation, home improvements, residential energy consumption.

* Corresponding author. Postal address: CERNA – Centre for industrial economics, MINES ParisTech, 60, boulevard St Michel, 75006, Paris, France. E-mail: matthieu.glachant@mines-paristech.fr

1. Introduction

The last report of the International Panel on Climate Change makes clear that climate is changing and global temperature will continue to increase even if greenhouse gas emissions are drastically cut (IPCC, 2013). Climate change will thus inevitably affect the economies. However, the costs of climate change remain largely uncertain, in particular because the knowledge on the capacity of human societies to adapt is limited[1]. In practice, adaptation consists of very different measures – the construction of dikes to protect from sea level rises, changes in crop-management practices in agriculture, the installation of insulation in houses to protect from heat, etc. – and different geographical areas face different challenges. Evaluating disaggregate costs of climate change at the sector, region, and country level is thus necessary to design appropriate policy support to adaptation and, more specifically, to calibrate the climate-damage function in Integrated Assessment Models (IAMs) which are widely used to forecast the long-term economic consequences of climate change (Nordhaus 2013).

This paper contributes to a growing empirical literature that produces ex post estimates of the impact of climate change (see the literature survey by Dell, Jones, & Olken, 2014). The usual empirical strategy exploits weather shocks within a given spatial area to identify the impact of climate change on various economic outcomes. This research stream – referred to as the new climate economy literature by Dell et al. – has examined the impacts of weather shocks on GDP, labour productivity, agriculture, conflict, crime, energy consumption, and health.

As explained by Dell, Jones, and Olken (2014), the main advantage of using panel data is identification. By including location-specific and time fixed effects, these studies better identify the causal impact of climate change than cross section or time series analyses that were conducted previously. However,

[1]For an overview of economic impact studies, see Tol (2009).

exploiting short-term variations in climate and economic outcomes makes it difficult to identify longer-term adaptation strategies, limiting the ability of these studies to forecast the future consequences of climate change. This weakness has been highlighted by Dell, Jones, and Olken (2014), Auffhammer & Mansur (2014) in their review of the studies on the impact on energy consumption, or Deschenes (2014) who has recently reviewed the empirical works on health-related impacts.

In this paper, we adopt a similar panel data approach, but we develop a new strategy to identify adaptation responses to temperature increases which we apply to the US residential sector. In order to present this strategy, it is first necessary to describe precisely what adaptation means in practice in this sector. When outdoor temperatures increase, home occupiers adjust their energy consumption in the short term. That is, they consume more electricity during heat waves if their home is equipped with air-conditioning; symmetrically, they reduce space heating during winters, and thus consume less gas or electricity depending on the heating technology installed. In the longer run, they also adjust the stock of durables installed in their dwellings: they can purchase new air-conditioners, change their heating equipment, or invest in weatherization (e.g. insulation, roofing and siding). More options are available if the time horizon is further extended: households will move in new dwellings that are more adapted to the new climatic regime (in particular, because they are located in less exposed areas); firms will develop new cooling and heating technologies, public authorities will redesign urban space, etc. Most of the existing empirical works only examine the short-term responses, holding fixed the stock of capital, the technologies and the institutional environment. In the specific case of the residential sector, Auffhammer and Aroonruengsawat (2011) and Deschênes and Greenstone

(2011)[2] propose forecasts of the impact of temperature increases on electricity consumption, but assuming no change in the stock of energy-related durables, no innovation, and no institutional change.

This paper does not analyse the entire sequence of adaptation responses, but we go one step further, by accounting for adaption investments. We use micro-data from 14 biannual and national waves of the American Housing Survey (AHS, 1985-2011), which includes information on energy consumption and the investments in weatherization, heating and cooling equipment performed in a large panel of US homes for around 150 localities in the US over the period 1985-2011. This data is matched with climatic data from the Global Historical Climatic Network (GHCN) Daily. We then use annual variations in location-specific temperature variables (cooling degree days, heating degree days) to identify the impact of temperature increases on the size of adaptation investments. The same data is used to estimate energy expenditures.

We then combine our econometric estimates with predicted changes in temperatures from a climate model to predict the impact of the A2 scenario on adaptation expenditures made by home occupiers in response to temperature increases (investment costs and energy costs). This scenario assumes a global average surface warming of 3.4°C in 2090-2099 relative to 1980-1999 (IPCC, 2007). The calculation of state- and month-specific temperature averages relies on the output of the Regional Climate Change Viewer (RCCV) which provides state-specific climate forecasts obtained by downscaling global climate simulations made with the ECHAM climate model.

Our best estimate of the present discounted value of the cost for adapting to temperature increases under the A2 scenario is $3,950 per housing unit, accounting both for energy and investment costs, but the null hypothesis of a zero cost cannot be rejected at conventional significance levels. To put this in

---

[2] Note that the main focus of this paper is to measure the welfare impacts associated with increases in mortality. In this context, they estimate electricity consumption in order to assess the economic cost of (short-term) adaptation.

context, this is equivalent to an annual expenditure of 210 $, which is approximatively 0.2% of the sample average annual household income or about 1.9% of the average purchase price of the housing units included in the sample (around 211,000 $). So this appears to be a small loss. The reason is that the installation and more intensive use of additional air-conditioners are partially offset by lower needs for space heating. Total residential energy expenditures would slightly increase by 5.1%. However, there would be a shift from gas to electricity: The model predicts an increase in electricity bills by 21% and a decrease in gas expenditure by 15%.

The empirical literature on adaptation in the residential sector is limited. As mentioned above, Deschênes and Greenstone (2011) and Auffhammer and Aroonruengsawat (2011, 2012). Auffhammer and Aroonruengsawat (2011, 2012) look at the impact of climate change on residential electricity consumption in California and find an increase in energy demand by around 3-6% under the A2 scenario (assuming no change in population and energy prices, and no adaptation). This result is interesting but cannot be generalised to the U.S: California has a very specific climate with little amounts of heating degree days and cooling degree days. Our estimates therefore provide a more general, U.S. level estimate. Deschênes and Greenstone estimate that, by the end of this century, residential energy consumption could rise by 11% in the US as a result of climate change. We go deeper into the analysis of the relationship between climate change and residential energy demand in several respects. First, we deal with both electricity and gas consumption. Extending the analysis to gas is crucial because it is widely used for space and water heating in the US. Furthermore, Mansur, Mendelsohn and Morrison (2008) analyse US energy demand in a setting in which fuel choice decisions are endogenous. Using a multinomial choice framework, they show that households prefer electricity to other fuels when temperatures are high, in particular because electric heating appliances have lower installation costs and fit more in regions where space heating is not intensive.

The remaining of this paper is structured as follows. The following section presents the conceptual framework. Section 3 describes the data and section 4 presents the estimation results. Section 5 assesses the magnitude of our estimates of the effect of climate change by simulating the A2 climate change scenario.

2. Analytical framework

2.1. Investment equations

In order to limit potential aggregation biases, we consider specific categories of investments that are related to adaptation. How far we can disaggregate the different types of adaptation investments is centrally determined by data availability. In this respect, the AHS data only allows us to construct two categories of adaptation-related home improvements: 1) the installation of major energy-consuming equipment, including major space heating appliances and air conditioners (either room or central air conditioners); and 2) weatherization, (i.e. addition/replacement of foam, weather stripping and caulking) and, by extension, improvements on doors and windows, roofing and siding that improve the energy integrity of dwellings. This means that we are not able to distinguish between space heating and air conditioning in the regressions.[3] But we will introduce assumptions in the simulations to go deeper in that direction.

Let I_{iht} denote the level of investment made by household i in category h (with h = equipment, weatherization) and in year t. To identify the relationship between this variable and temperature variations, we fit the following linear equation:

$$I_{iht} = \alpha_h Z_{it} + \beta_h X_{it} + \mu_{ih} + \tau_{ht} + \varepsilon_{iht} \qquad (1)$$

[3] Another category of lower interest corresponds to all the other indoor investments that are not directly related to climate change: changes to the bathroom; changes to the kitchen; home extensions; and any other major indoor improvement. These are studied in Online Appendix B.6 as a means to assess overall robustness of results.

where Z_{it} is a vector of climate variables that will be described in the next section. Importantly, these variables vary across locations and time, and thus across dwellings. X_{iht} is a vector of control variables such as household size. The choice of these control variables is discussed in detail later on. Finally, μ_{ih} is a household-category fixed effect; τ_{ht} is a time fixed effect and ε_{iht} is the random error term. α_h and β_h are category-specific parameters to be estimated.

A usual concern in the empirical literature is that investment are lumpy with long periods with no investment ($I_{iht} = 0$) interrupted by more active investment periods (e.g. Doms and Dunne, 1998). Dixit and Pindyck (1994) provide theoretical foundations for such behaviour. In our case, households may prefer to make all the necessary improvements at one point in time because of the hidden fixed costs. In particular, home renovation limits the ability to live in the house while it is being renovated.

A popular approach used to deal with this problem interprets home improvements as a left-censored variable, assuming investments are only observed with a positive value. In Online Appendix B.1, we estimate such a latent variable model. The model is similar to that of Helms (2003), except that we fully take advantage of the panel structure of the data. The signs and relative magnitude of the coefficients appear to be very similar to those obtained with the linear model. We however keep the linear model as our base specification because it produces estimates of the fixed effects and is therefore more convenient when it comes to making predictions.

Climate variables

Choosing the right climate variables to be included in the vector Z_{it} is obviously critical. They should be closely related to the adaptation of dwellings to temperature changes. We primarily use heating degree days and cooling degree days. These are usual measurements designed to reflect the demand for heating and the demand for cooling, the two housing services most directly affected by global warming[4]. The precise definition of cooling degree days is the number of degrees that a day's average temperature is above 65° F when it is assumed that people start to use air conditioning to cool their buildings. Symmetrically, heating degree days is the number of degrees that a day's average is below 65° F.

A drawback of these two indicators is that they fail to account for potential non-linearities in the marginal impact of temperature change on investments. That is, it is assumed that a one degree increase has the same effect on investment whether it occurs in a mild day when the temperature is 70°F or during a heatwave where the temperature is 90°F. As a robustness checks, we thus also estimate a more flexible specification including temperature bins which gives estimates of the specific impact of different ranges of temperatures on investments (like in Deschênes and Greenstone, 2011). Results are not substantially different. In addition to temperature variables, we control for the confounding effect of precipitations by including annual precipitation levels in the equation.

A further issue is that investments are not determined by the *contemporaneous* value of these variables. As the lifetime of investments in housing is relatively long and the benefits from installing air conditioning or insulating in a house depend on future needs, households should choose their investment based on their expectations about future weather, not on the current temperatures. Moreover, even without assuming that households look forward, they are aware of climate inter-annual variability, meaning that their view of the current climate is based on some averaging of past climate values.

This has led us to adopt an adaptive expectation framework whereby consumers adjust their expectations based on past experience.[5] Specifically,

---

[4] See: http://www.degreedays.net/introduction.

[5] Gelain and Lansing (2014) provide recent evidence that backward-looking expectations may operate on the housing market. They argue that such expectations are a better predictor of high volatility in price-rent ratios compared to rational expectations.

assume that Z_{it} is actually consumer *expectation*. The adaptive model assumes the following relationship:

$$Z_{it} = z_{it-1} + \lambda(z_{it-1} - z_{it}) \tag{2}$$

Where z_{it} is the real (realized) value of the climate variables at year t. Eq. 2 means that current expectations are composed of past expectations and an "error adjustment" term, which raises or lowers the expectations depending on the realized value of Z_{it}. The parameter $\lambda \in (0,1)$ captures the adjustment speed between past and current expectations. Applying Eq. (2) recurrently over all past periods, expectations at time t of the temperatures at t + 1 are equivalent to an exponentially-weighted moving average:

$$Z_{it} = \lambda \sum_{k=0} (1-\lambda)^k \left(z_{it-k-1}\right)^{k+1} \tag{3}$$

The value of λ is not directly observed. We therefore make the assumption that all households use a value that would make their predictions as accurate as possible. Empirically, we estimate λ with a non-linear regression based on Eq. (3)[6]. The regression aims to predict current weather based on a weighted average of past weather for all households in our data. Taking the sum of cooling and heating degree days, we obtain $\lambda \approx 0.31$. This is equivalent to assuming that expectations primarily rely on the past 7-8 years.[7]

Control variables

Choosing the adequate controls is complicated by the fact that the climate potentially influences variables that are possible candidates. Take the example of household income. It is an obvious candidate control as it influences the

[6] This equation is slightly modified to account for the fact that we have a limited amount of lags in the model (l =15). We assume that $Z^*_{it} = \left(\lambda \sum_{l=0}(1-\lambda)^l \, Z^{l+1}_{it-l-1}\right) / \left(\lambda \sum_{l=0}(1-\lambda)^l\right)$ since $\lambda \sum_{l=0}(1-\lambda)^l$ tends to one only when l tends to infinity for low values of λ.

[7] We obtain similar results with a distributed lag model that includes three year lags as control variables. In contrast, results are significantly different when using current values z_{it}, that is, when assuming that households (irrationally) only consider contemporaneous heating degree days and cooling degree days when making long-term decisions (Appendix A.1).

propensity to invest. But it is also reasonable to assume that temperature has an impact on its level. This has actually been shown by recent empirical studies (e.g., Dell et al. 2009). If this variable is included in the equation, the coefficients α_h will then not identify the full effect of climate. It will only capture the direct effect of temperature, ignoring the indirect effect that passes through changes in income. This will then skew the results of our simulations (except if we have estimates of the impact of the temperature on income to calibrate the simulation model). To reduce the risk of "over-controlling", Dell et al. (2014) suggest controlling for factors that are assumedly not influenced by the climate.

Accordingly, in addition to income, we do not control for several other variables impacting investments. We do not incorporate information on the local price of each type of investment *h*, because the correlation between temperature and local prices for energy-related investments is likely to be determined by temperature: e.g. a higher demand for air conditioning during very hot summers will increase the local price of air-conditioners. In the same vein, we do not control for electricity and gas prices, although they are clear determinants of the demand for heating and air conditioning equipment: These prices are obviously influenced by local climate factors since temperature has direct impacts energy production, transmission and distribution. Finally, we do not control for the impact of past investments on current investments, because past investments depend on past expectations about the climate. Hence, they are correlated with current climate shocks in a causal manner, provided that expectations are formed with some rationality.

In the end, we limit ourselves to including time and household fixed effects, the number of individuals living in the house and connection to pipe gas as controls. As a result, the coefficients α_h do not only identify the direct causal effect of a change in temperature on the demand for investments, but the correlation between temperatures and investment levels, including income-effects, demand-side and supply-side effects. This assessment of temperature changes then yields a more complete measurement of their overall impact on investments.

A last concern pertains to the existence of energy efficiency policies in certain States (e.g., tax credits or subsidized loans) which are actually implemented to influence investments in space heating, air cooling or weatherization. These policies are likely to be correlated with climate shocks. We could still want to control for them if we think future policies are unlikely to resemble current policies. In this respect, we simply exclude from the sample all the observations in which households have benefited from energy efficiency subsidies (2% of the observations). As a robustness check, we use this piece of information to construct a dummy variable which is included in the equation. As this variable is likely to be endogenous – it reflects the existence of policies promoting energy efficiency at a local level[8] –, we use a control function approach. There is very little difference in the results obtained by the two approaches (see the online Appendix B.2).

2.2. Energy expenditures

We need to estimate energy expenditures for two reasons: because adjusting energy consumption the short-term strategy to adapt to temperature variations regardless of the investments that are made, but also because adaptation investments modify the level of energy use (downwards – through weatherization – or upwards – through air conditioning).

We separately estimate the demand for electricity and gas, which are the two main energy sources used in the US residential sector. We choose a dynamic

[8] In particular, this variable only captures information about those households that actually performed alterations. For the other households, we do not know if they could have had access to government support or not. In addition, this is a binary variable whereas household choices are driven by the size of subsidies.

model with a log-log functional form which is very similar to that used by Alberini et al. (2011) except that we account for the impact of investments. The use of a dynamic model is justified by the fact that energy consumption slowly adjusts over time due to persistent consumption patterns and habits within a household.

The dependent variable is E_{ift}, the logarithm of the annual energy expenditure in fuel *f* (with f = gas, electricity) of household *i* at time *t*:

$$\ln(E_{ift}) = \gamma_f \ln(E_{ift-1}) + \theta_f \ln(z_{it}) + \sum_{h=1}^{3} \phi_{hf} \ln(K_{iht}) + \omega_f Y_{it} + \mu_{if} + \tau_{ft} + \epsilon_{ift} \quad (4)$$

In this equation, z_{it} is the vector of climate variables, which includes the heating degree days, the cooling degree days and the number of days with precipitations. In contrast with the investment equations, we use contemporaneous values and not the expected values since energy consumption is only influenced by contemporaneous temperatures.

K_{iht} measures the amount of capital in the housing unit in each category h. It is equal to the discounted flow of past and current investments (I_{iht}, I_{iht-1}, I_{iht-2}…). The precise calculation is described in Appendix A.5[9]. Note that we consider three investment categories: the installation of major energy using equipment and weatherization that are estimated in the preceding equations, but also a third category including all the other investments. The reason is that they could also influence of energy consumption. The inclusion of E_{ift} is what most distinguishes this paper from previous works on the impact of climate on residential energy use (for a review, see Auffhammer and Mansur, 2012),

In Eq.(4), Y_{it} is the vector of controls which includes the log of family size and connection to pipe gas. For the reasons discussed previously about investment

[9] One difficulty is that the investments are not observed before the year of purchase or construction. We estimate the stocks in that year relying on the sales price. All the details are provided in Appendix A.5.

equations, we do not control for income and the energy prices as they are likely to be influenced by temperature. The equation includes a full set of household-by-fuel fixed effects, μ_{if}, which absorb all household-specific time invariant household specificities. It also includes a set of time-by-fuel dummies, τ_{tf}, which, for instance, control for the general evolution of energy prices that might affect energy consumption. ϵ_{ift} is an error term. Finally, γ_f, ϕ_f, θ_f, and ω_f are (vector of) parameters to be estimated.

The fact that Eq. (4) is estimated separately for gas and electricity potentially generates a sample selection bias as households may select the type of fuel used in their homes. The risk is however limited. The equation includes household-by-fuel effects; fuel selection prior to moving into the house is thus controlled for. We also control for the availability of piped gas, a major reason for choosing gas over other types of fuels. More generally, fuel switching, which mostly concerns space heating as air conditioners use electricity, occurs when installing a new equipment. And, this is infrequent even in this case. In the AHS data, households report a change in main heating fuel concomitant to home improvements in equipment in only 0.6% of the observations.

Endogeneity

Estimating Eq. (4) requires dealing with the endogeneity of several variables: the lagged dependent variable, and the different stocks of capital.

To find an exogenous instrument for the first variable is particularly difficult as the endogenous variable and the dependent variable are tightly related. A standard strategy in dynamic panel data models is to use deeper lags. This approach requires the arguably strong assumption that there is no complex serial correlation structure in the dependent variable. To avoid making this assumption, we opt for a different type of instruments which captures exogenous variations in the time spent in the house. A first instrumental variable is the log of average commuting time from home to work in year $t-1$ of all the members of the household over 14. Commuting time is necessarily correlated with the time spent in the house – it indicates less leisure time – and thus with the consumption of energy-using housing services. We also believe that commuting time at $t-1$ is not correlated with the error term of Eq (4) at t. The validity argument is that there is no persistent direct effect of commuting time on energy consumption. A potential problem is that commuting time at $t-1$ is correlated with commuting time at t because both are driven by long term choices of where to live, work or study. And commuting time is obviously contemporaneously correlated with variables like income or energy prices that are included in the error term. To avoid the problem, the (log of) commuting time at t is included in Eq. (4) as an additional control variable. Interestingly, commuting time also depends on multiple factors that are clearly uncorrelated with energy consumption, such as changes of school, in departure time, in road traffic conditions, or in transit availability.

The impact of commuting time on energy consumption is stronger if the house or the flat is big (simply because energy consumption depends on dwelling size). In order to strengthen the instrumentation, we thus add a second instrumental variable which interacts commuting time at $t-1$ with dwelling size (in square feet and in logs) at the time of purchase[10]. As a result, commuting times can have a different effect on energy expenditure depending on the size of the house. We also add the same variable but at time t as a control.

Next, the different stocks of capital in the different types of investments could be simultaneously determined by factors not already included in the model. This may occur because the current stocks of capital include the investments made at time *t* and we use observed values for K_{iht} and not those predicted by the home improvement model. The use of observed values is preferred since

[10] We use the size of the house at the time of purchase and not at time t to avoid our instrument to capture the effect of home extensions (i.e. between time t-1 and time t) on energy expenditure.

the investment model tends to predict changes in long-term averages relatively well but is not efficient when it comes to predicting the likeliness that specific investments happen at time t and not t+1 or t-1. To deal with the endogeneity of K_{iht}, we instrument all the stocks of capital in every category h with their lagged values.

The validity of all the instruments is corroborated by Hansen tests run concomitantly to the Blundell-Bond estimator. They pass the tests for all the specifications that are estimated.

Estimation method

We estimate Eq. (4) using a system-GMM estimator (Blundell-Bond) to extract as much information as possible from the instruments and ensure instrument strength. By preferring the system-GMM estimator to the Arellano-Bond estimator, we impose an additional assumption: the mean stationarity in the dependent variable. This assumption seems reasonable in the case of residential gas and electric demand in the U.S. since the real value of residential energy bills has not evolved sharply for the period that is studied (see figure below). However, we cannot test for mean stationarity using unit root tests since our panel is not balanced.

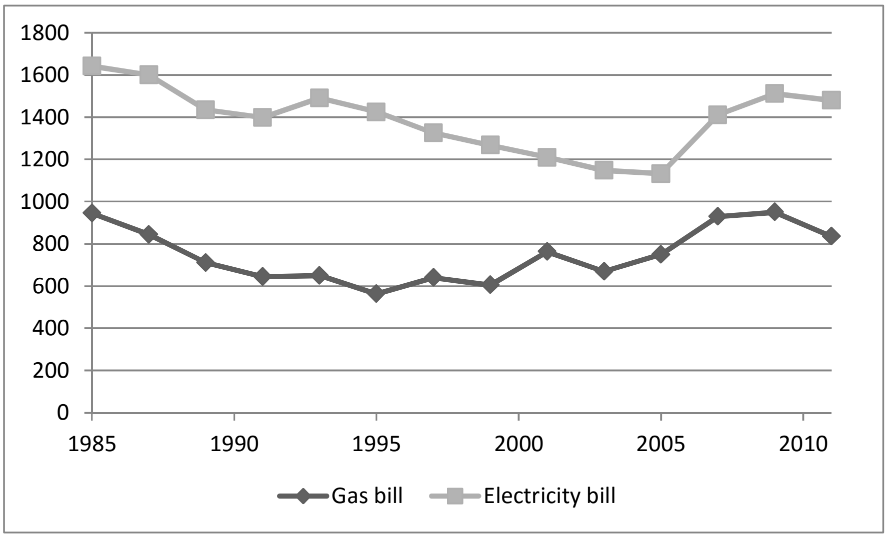


Figure 1: Average energy bills in the sample used (in real 2011 dollars)

To check the robustness, we report results of alternative models in Online Appendix B.X. A first alternative would be a simple fixed effects model. It would however suffer from an omitted-variable bias as it would not include the lagged dependent variable. To bypass this obstacle, we use instead a fixed effect model with an AR1 disturbance autoregressive term. This model is known to be in-between dynamic panel data models and a fixed effect models, but do not tolerate endogenous variables. The second alternative is to use the Arellano-Bond estimator. Results are provided in Online Appendix B.5 for the Arellano-Bond estimator and online Appendix B.6 for the fixed effect model with an AR1 disturbance autoregressive term. The results of the Arellano-Bond estimator are similar to those of the fixed effect model, suggesting weak instrumentation in the case of the Arellano-Bond estimator. Most of the coefficients estimated with the fixed effect model are similar to the coefficients obtained with the Blundell-Bond estimator. However, the fixed effect model finds an unrealistic, positive correlation between gas demand and cooling degree days, and no correlation with heating degree days. This suggests an omitted variable bias due to the exclusion of the lagged dependent variable. Using a fixed effect model with an AR1 disturbance term appears to correct for this bias to a large extent, suggesting superiority of a dynamic model over a fixed effect model.

3. Data

We principally rely on three main data sources: the American Housing Survey for data on housing units, home improvements, energy consumptions and household characteristics, the Global Historical Climatology Network (GHCN) Daily for meteorological data and the ECHAM model for climate change predictions. This section briefly describes the data sources and reports summary statistics.

3.1. Data sources

The AHS data

The data on housing units, home improvements, energy consumptions and households are essentially taken from the national sample of the American Housing Survey which covers Metropolitan Statistical Areas (MSAs). We use longitudinal data from 14 waves of the national AHS from 1985 to 2011.[11] The study relies on data from the 126 MSAs with more than 100,000 inhabitants. Information on the location of each housing unit – a decisive variable to relate each unit to local climatic conditions – is not available for smaller MSAs. These large MSAs are spread all over the United States, thereby experiencing very different climatic conditions.

The AHS includes information on nine different types of home improvements. The weatherization variable is constructed by adding the investments made in the following four categories: roofing; insulation; siding; and storm doors and windows. The equipment variable is identified as a single category in the survey. As mentioned previously, the AHS does not unfortunately distinguish investments in air conditioning from investments in space heating[12]. Nevertheless, households report their main heating fuel and their main fuel for air conditioning. This information is used in some specifications of the energy consumption model, to construct interaction variables used to roughly infer the specific impact of investments in heating and cooling equipment (more detail is provided in section 5). Note also that this data only describes owner-occupied units for which the information on renovation investments is available.

[11] The waves before 1985 cannot be used in a panel data analysis because the AHS underwent a redesign in 1985 and the units observed before 1985 and after 1985 are different.

[12] In 1997, the typology was refined but we had to stick to the previous typology to be able to use the entire study sample from 1985.

For each type of home improvements, we observe the level of annual investments between 1985 and 2011. The value of the stock of capital already embodied in a home before 1985, a needed piece of information to construct the capital stock variables included in energy equations, is however missing. We derive this initial stock – and from this the value of K_{iht}– from the purchase price or the construction cost of the housing units as registered in the American Housing Survey after a transaction or after construction for new buildings. The exact method is presented in Appendix A.5.

The AHS also provides information on home occupiers. In particular, it identifies the moment when a household left and when a new one moved in a given housing unit. This information is used to construct household-specific fixed effects. Information on the level of energy expenditures, on whether the neighbourhood has access to pipe gas, and on commuting times is also extracted from the AHS.

Note that the precise location of each housing unit is not available, which leads us to take the centroid of the MSA as a proxy. This is not a major problem as temperature change patterns are unlikely to vary much within a given MSA. This would be more problematic if we examined the impact of other more spatially-differentiated climatic events such as floods and hurricanes.

Weather data

The weather data is drawn from the Global Historical Climatology Network (GHCN) Daily. We extracted land-based (*in situ*) historical observations recorded from 1970 to 2011 by 22,000 meteorological stations that match the MSAs included in our sample and that operate a minimum number of days over the year. More specifically, we kept all the stations within a 50km radius of the centroid of an MSA and which has daily information on at least 20 days in each month of the year and calculate averages for each MSA.

The key variables are the annual heating degree days, cooling degree days, and the annual precipitations in millimetres. We however run alternative specifications using the number of days that fall within 10°F temperature bins (the first bin is “below 10°F” and the last one “above 90°F”). We also compute the number of days with precipitations above 50mm, 100mm and 200mm. The use of these variables offers a more flexible estimation of the impact of temperatures and precipitations on home improvements and energy demand.

Climate change prediction data

State-level monthly average temperature predictions are drawn from the 5th version of ECHAM, an atmospheric general circulation model developed at the Max Planck Institute for Meteorology. We focus on the A2 scenario and its predictions for the end of the century (2080-2099). A2 is a business-as-usual scenario in which a relatively high amount of GHG emissions is released into the atmosphere, leading to a global average surface warming of 6.1°F in 2090-2099 relative to 1980-1999 (IPCC, 2007).

US State-specific averages are accessible with the Regional Climate Change Viewer (RCCV) of the U.S. Geological Survey. The RCCV uses a downscaling method of the output of ECHAM, averages temperatures within States, and then compares the historical period of 1980-1999 with the ECHAM model output for 2080-2099. This gives a predicted daily mean temperature increase for each month and State. This increase is added to all the days of the historic weather data to compute daily average temperature forecasts for 2080-2099. We then use these daily temperature forecasts to predict State-level changes in heating degree days and cooling degree days.

3.2. Summary statistics

Investment, energy expenditure, and household data

Table 1 provides the list of the descriptive statistics from the AHS data that are used for the model estimation. The sample is composed of 58,529 observations[13] from 126 Metropolitan Statistical Areas[14]. The investment frequency is quite low (7.4% for the installation of equipment and 16.6% for weatherization), but the average investment is significant ($4,000 and $5,000, respectively). This lumpiness could justify the use of a latent variable model; tobit results displayed in Appendix A1 are however similar to those of the base linear model. Adaptation-related investments do not constitute the majority part of renovation expenditures. In particular, the capitalized investment in equipment is minor compared to the other categories.

[13] This is much less than the 262,872 observations of geographically-located and owner-occupied units between 1985 and 2011. But many values are missing, in particular the values of the purchase price or the construction cost. Outliers have also been excluded (see more details in Appendix XXX).

[14]

Table 1: Descriptive statistics of the AHS data

Variable	Unit	Mean	Std. deviation
Investments in equipment			
Capitalized investments	$	10,201	7,641
Respondents declaring an investment	%	7.4	-
Expenditure if an investment is made	$	3,978	2,891
Investments in weatherization			
Capitalized investments	$	54,534	40,732
Respondents declaring an investment	%	16.6	-
Expenditure if an investment is made	$	4,817	4,904
Investments in other indoor amenities			
Capitalized investments	$	104,368	77,114
Respondents declaring an investment	%	29.1	
Expenditure if an investment is made	$	6,730	10,101
Energy expenditure and consumption			
Annual electricity expenditure	$	1,379	819
Annual gas expenditure	$	742	723
Annual electricity consumption	MM.btu/year	36.7	23.0
Annual gas consumption	MM.btu/year	64.8	63.0
Other relevant variables			
Number of people in household	#	2.82	1.52
Housing units connected to pipe gas	%	79.1	-
Commuting time	min.	22	16
Square footage of unit	sq. ft.	2,189	1,267
House price at time of purchase	$	211,310	174,966

Notes. Source: AHS. Survey years: 1985-2011. Max. number of observations: 58,529. Comments: all the variables in dollars are expressed in 2011 real dollars. The correction of nominal values has been made using the U.S. Consumer Price Index of the Bureau of Statistics of the U.S. Department of Labour.

Weather and Climate Change Statistics

Detailed weather statistics for the entire sample and by U.S. climatic region are provided in Table 2. We report information on daily temperature, number of heating and cooling degree days and the number of days below 10°F and above 90°F. Using the same format, Panel B presents the impacts of climate change based on the ECHAM model and for the A2 scenario. These figures show a high heterogeneity between regions in daily temperature, but also in the number of days with extreme temperature (cold or hot). The USA is obviously not representative of the climatic conditions observed all over the world level, but they nevertheless provide a significantly diversified sample.

Figure 2 below plots the average number of days falling within a given temperature bin. The grey bars report the averages as observed with the GHCN data of NOAA over the 19985-2011 period and the black bars report the averages as predicted by ECHAM under the A2 scenario. The A2 scenario predicts a dramatic increase in hot days (80-90°F) and very hot days (>90°F) in 2090-2099, mostly in hot regions. In contrast, the number of days below 70°F decreases rather uniformly across the different temperature bins.

Table 2: Summary statistics of climate data

Annual averages	Daily temperature	Heating degree days	Cooling degree days	Days <10°F	Days > 90°F
Panel A. Historical temperature data (1985-2011)					
All housing units	57.8	4,077	1,441	2.6	2.5
Cold regions	51.3	5,882	863	5.1	0.0
Central	52.2	5,663	979	4.7	0.0
West North Central	50.4	6,343	1,009	11.4	0.0
Northwest	48.6	6,112	116	0.4	0.0
East North Central	47.6	7,023	670	12.5	0.0
Northeast	52.7	5,404	904	1.2	0.1
Hot regions	64.4	2242	2033	0.2	5.0
South	65.0	2474	2479	0.5	2.1
Southeast	67.0	1891	2612	0.0	0.0
Southwest	61.8	3479	2326	1.0	29.6
West	62.8	1982	1194	0.0	2.9
Panel B. Predicted change from ECHAM model under the A2 scenario (2080-2099)					
All housing units	+7.5	-1425	+1,317	-16.2	+16.2
All housing units	+7.5	-1,411	+1,326	-1.8	+17.0
Cold regions	+7.7	-1,846	+977	-3.5	+3.2
Central	+7.5	-1,718	+1,040	-3.4	+4.4
West North Central	+7.4	-1,684	+1,026	-6.2	+7.1
Northwest	+6.3	-1,825	+485	-0.3	+0.0
East North Central	+7.7	-1,959	+848	-8.0	+1.5
Northeast	+8.0	-1,898	+1,018	-1.1	+3.2
Hot regions	+7.2	-963	+1,682	-0.2	+31.1
South	+7.8	-921	+1,913	-0.3	+64.6
Southeast	+6.9	-784	+1,727	0.0	+28.4
Southwest	+8.2	-1,329	+1,670	-0.7	+36.6
West	+6.9	-1,014	+1,506	0.0	+10.8

The climate variables are averaged over all the observations of the AHS datasets used in the regressions. Hence, regional averages are not representative of the regions but of the sample of housing units within each region. When a unit is located in a Metropolitan area which overlaps two or three states, it enters the calculation of the averages in all the States it overlaps, but with a weight of 1/2 or 1/3.

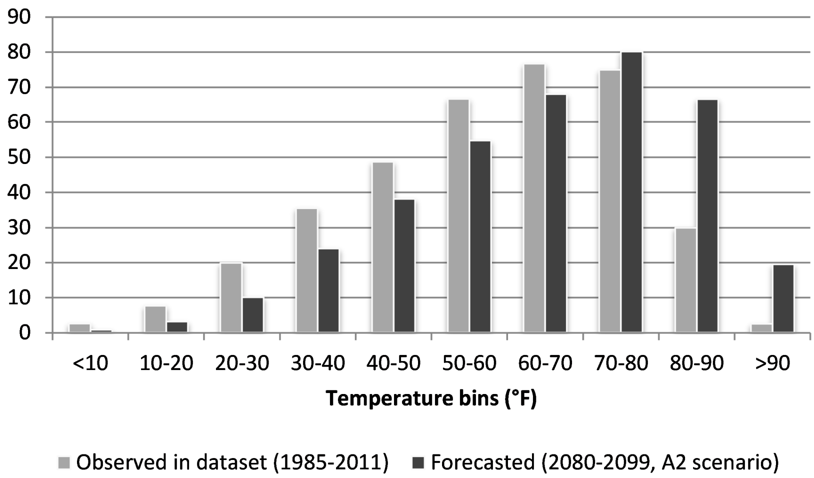


Figure 2: Observed and forecasted number of days falling within each temperature bin

The number of days is averaged for all the observations in the AHS datasets used in the regressions.

4. Results

This section is divided into two subsections. The first provides estimates of the relationship between temperature increases and the level of adaptation investments, and the second examines the impact of temperatures and investments on the size of energy expenditures.

4.1. Adaptation investments

The base results for investments in equipment and in weatherization are displayed in Table 4. Heating and cooling degree days are used as the reference climate variables. In Appendices A.1 to A.3 and Online Appendices B.1 to B.6, most of the hypotheses used to calibrate the models are tested in a series of robustness checks which confirm our findings.

We find a statistically significant impact of expected heating and cooling degree days on investments in equipment. The two coefficients are positive, consistent with the interpretation that households purchase more (or larger) heaters when winter temperatures fall and more air conditioners when summer temperatures increase.

The impact of heating degree days on weatherization is also positive and significant, but the results show insignificant impact of cooling degree days. It may be that households prefer air conditioners over insulation to adapt to high temperatures for cost reasons: compared to weatherization, air conditioning generates variable costs which are limited if equipment only operates the low number of days when the temperature is very high. Note that precipitations have unsurprisingly no significant impact.

Table 3: Main results for investments in energy-related home improvements

Type of investment	Equipment	Weatherization
Expected heating degree days	0.161^{**}	0.322^{**}
	(2.23)	(2.08)
Expected cooling degree days	0.354^{***}	0.297
	(2.69)	(1.13)
Expected precipitations	-0.00399	0.0141
	(-0.39)	(0.58)
No. people in unit	-4.347	41.11^{*}
	(-0.42)	(1.70)
Connection to pipe gas	89.83	138.1
	(1.55)	(1.43)
Observations	44,975	42,900

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported.

Table 4 displays the results of alternative models using temperature bins. The existence of non-linearities in the relationships between heating and cooling degree days on the one hand and investments on the other hand is confirmed. The impact at the extremes is stronger than that observed in a comfort zone around 60-70°F.

Table 4: Linear investment models using temperature bins

Type of investment	Equipment	Weatherization
Expected # days with temperature:		
Below 10°F	1.080	16.10
	(0.16)	(0.99)
Between 10-20°F	12.27**	25.51*
	(2.00)	(1.92)
Between 20-30°F	-0.775	4.304
	(-0.17)	(0.44)
Between 30-40°F	7.682**	11.67
	(2.31)	(1.62)
Between 40-50°F	1.399	0.501
	(0.49)	(0.08)
Between 50-60°F	1.897	5.448
	(0.82)	(1.05)
Between 60-70°F	-	-
Between 70-80°F	0.610	2.424
	(0.31)	(0.54)
Between 80-90°F	5.977**	2.808
	(2.07)	(0.48)
Above 90°F	16.80**	19.73*
	(2.43)	(1.80)
Expected days with precipitations:		
No precipitation	-	-
Between 0-50mm	0.596	0.980
	(0.60)	(0.44)
Between 50-100mm	-4.553	-6.036
	(-1.26)	(-0.87)
Between 100-200mm	1.398	6.974
	(0.36)	(0.87)
Above 200mm	0.0169	0.642
	(0.00)	(0.06)
No. people in unit	-4.294	40.93*
	(-0.42)	(1.70)
Connection to pipe gas	91.51	139.1
	(1.58)	(1.44)
Observations	44,975	42,900

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies.

4.2. Energy consumption

Table 5 displays results for gas and electricity consumption. For each fuel, we estimate the base model described by Eq.(4) and a variant where we interact the equipment capital with the fuel used to heat and cool the house. Models (1) and (3) thus estimate the average impact of a change in equipment on energy demand across all households and years while this impact can differ according to the main heating and the main cooling fuel declared by households with the models (2) and (4). The nature of the fuel provides a direct indication of the type of equipment installed in the housing unit. Interacting with the equipment investment variable will help mitigating the problem of AHS data that does not distinguish between investments in heating or cooling devices when running simulations in the next section.

The coefficients of the control variables have the expected signs: past consumption has a persistent impact on current consumption ($\gamma_f \approx 0.53$ for electricity and about 0.30 for gas), family size drives consumption upwards and connection to pipe gas encourages households to choose gas heating.

In the base specifications (1) and (3), households use more gas when the number of heating degree days increases, but the impact on electricity consumption is not significant. This difference is consistent with the fact that only 21% of the households declare electricity as their main heating fuel. That cooling degree days increase electricity consumption is also intuitive as most homes are equipped with electric air conditioners.

In contrast, the significant and positive impact of cooling degree days have on gas expenditures is less easy to interpret as gas is scarcely used for cooling (around 5% of the households in the sample). This result is however interesting because it illustrates that changes in energy prices are a channel of influence of climate change on adaptation expenditures. In the US, use of natural gas has two seasonal peaks. The first peak occurs during the winter, when cold weather increases the demand for natural gas space heating in the residential and commercial sectors. A second peak occurs in the summer when air conditioning use increases demand for electric power, an increasing portion of which is provided by natural gas-fired generators. Surprisingly enough and for reasons related to the gas storage technologies, gas prices only peak in summer months. The seasonal variation is high: during the study period (1985-2011), the annual maximum price of natural gas delivered to residential consumers, generally observed either in July or in August, was in average 45% higher than the minimum price observed in December or January.[15] In Figure 3, this pattern is illustrated by plotting the evolution of residential gas prices between January 2005 and December 2011. We can now interpret the results of model (4). Households use gas in summer months to fuel water heaters, cook stoves, dryers, and other equipment, implying that the summer price peak significantly inflates their energy bill. The global warming expected in the future would reinforce this effect. Increasing electricity consumption of air conditioners, higher summer temperatures would boost gas consumption by electric power generators, and thus gas prices in summer months.

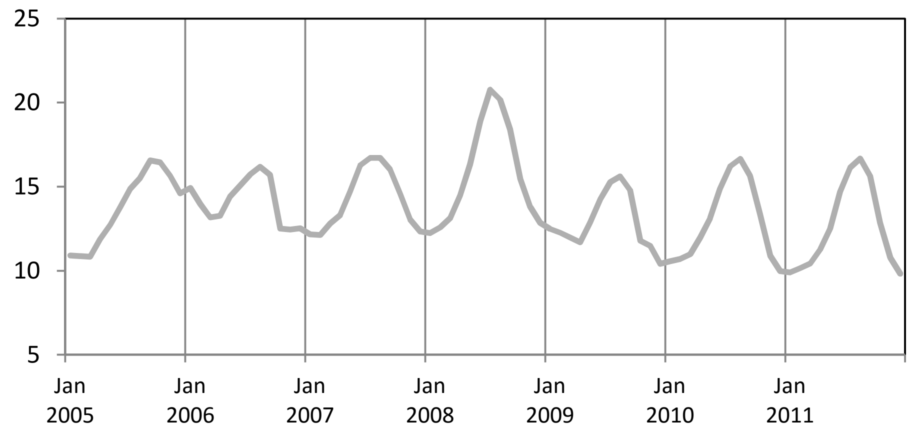
[15] See: https://www.eia.gov/dnav/ng/hist/n3010us3m.htm

Table 5: Blundell-Bond estimation of energy expenditure models

Type of fuel	Electricity		Gas	
	(1)	(2)	(3)	(4)
Lagged dependent variable	0.558***	0.571***	0.359**	0.262*
	(5.35)	(5.75)	(2.35)	(1.67)
Heating degree days	-0.00201	0.00231	0.257***	0.266***
	(-0.34)	(0.39)	(4.29)	(4.67)
Cooling degree days	0.0876***	0.0646***	0.0475***	0.0534***
	(3.13)	(2.87)	(3.47)	(3.83)
Capital in equipment	0.00834	-0.00644	0.0408**	0.0130
	(0.61)	(-0.48)	(2.06)	(0.67)
x heating fuel is electricity		0.00827***		
		(5.16)		
x AC fuel is electricity		0.00619***		
		(3.49)		
x heating fuel is gas				0.0337***
				(5.52)
x AC fuel is gas				0.00345
				(1.33)
Capital in weatherization	0.0193	0.0249	-0.0471	-0.0636*
	(0.85)	(1.09)	(-1.43)	(-1.86)
Capital in other amenities	0.0240	0.0269	0.0672*	0.0882**
	(0.92)	(1.02)	(1.89)	(2.36)
Precipitations	0.0555***	0.0511***	0.0890***	0.118***
	(3.61)	(3.62)	(3.52)	(4.17)
Connection to pipe gas	-0.105***	-0.0616***	0.149***	0.0726
	(-4.14)	(-3.01)	(2.66)	(1.23)
No. people in unit	0.131***	0.130***	0.0917***	0.110***
	(5.90)	(6.08)	(4.68)	(5.19)
Average commuting time	-0.0617***	-0.0536***	-0.0819***	-0.0891***
	(-3.30)	(-3.22)	(-3.43)	(-3.75)
x sq. footage at time of Purchase	0.0648***	0.0587***	0.0816***	0.0921***
	(4.00)	(4.03)	(3.38)	(3.77)
Observations	19,598	19,598	13,718	13,718
Hansen test	0.24	0.26	0.16	0.22

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. Lagged dependent variable is instrumented using first lag of average commuting time and interaction between commuting time and square footage of unit at time of purchase. All (dependent, independent and instrumental) variables are in logarithm.

Figure 3: U.S. price of natural Gas delivered to residential consumers from Jan 2005 to Dec 2011 (in dollars per cubic feet)



Turning next to the capital variables, the stock of equipment appears to have a positive impact on gas expenditures in model (3). When the variable is interacted in model (4), results are similar. Equipment investments are strongly correlated with consumption when gas is the main heating fuel; a positive effect is also recorded when gas is used for air conditioning, but the latter is very small, plausibly since air conditioners that use gas as a fuel are rare. The impact on electricity expenditures is not significant in model (1), but becomes positive and significant when we account for the type of fuel used by households (model 2).

The effects of weatherization on energy expenditures tend to be less precisely estimated as coefficients are not statistically for electricity. But it shows the expected negative sign for gas in model (4). Note also that capital in other amenities is positively correlated with gas use in this model, which indicates that some of these investments increase heating needs, e.g. in the case of an home extension.

Similar results are obtained when slight changes to the specification are performed. In Appendix A.4, we use alternative climate variables. The results are analogous: hot days lead to a higher consumption of electricity and cold days to a higher consumption of gas. In Online Appendix B.7, we run the econometric model using energy consumption instead of energy expenditure and control for electricity and gas prices, which are instrumented due to endogeneity. The magnitude and statistical significance of the coefficients are very similar.

5. Impacts of climate change

This section aims to exploit the empirical results on investment and energy consumption to estimate the impact of higher temperatures on energy consumption and the resulting adaptation cost. The predictions are calibrated for the A2 scenario of IPCC and they rely on the ECHAM model of the Max Plank Institute, as described in section 3.

5.1. The simulation procedure

The simulation procedure involves three steps. First, we use the estimated coefficients of the fixed effect investment models to compute the impact of a change in expected heating or cooling degree days on the average amount of equipment and the weatherization level in the house. For example, the estimated impact of predicted temperature changes in equipment investments in a given housing unit and year is calculated as follows:

$$\Delta I_{iet} = \hat{\alpha}_{e1}\Delta HDD_{it} + \hat{\alpha}_{e2}\Delta CDD_{it}$$

That is, the predicted change in heating and cooling degree days ΔHDD_{it} and ΔCDD_{it} is multiplied by the corresponding impact on investment ($\hat{\alpha}_{e1}$ and $\hat{\alpha}_{e2}$). These numbers are then averaged over all housing units and years to give an estimate of the average change in the level of equipment investments. The same calculation is performed for weatherization investments.

In a second step, we use the output of the dynamic panel models of energy expenditure to derive energy expenditure estimates that account for potential adjustments in capital calculated in the first step. This calculation requires introducing several assumptions. First energy expenditure model provides information on: 1) the impact of equipment on energy expenditure; 2) the impact of equipment when fuel f is used for heating; and 3) the impact of equipment when fuel f is used for cooling. We make an efficient use of these parameters (which we match with our three variables for changes in equipment) to make a more accurate prediction of energy expenditure levels while relaxing the strong assumption that the relative shares of capital for heating and for cooling should remain constant.

Second, while investment models predict an average investment value, the energy consumption model requires the total capital stock to be computed. We simply assume that the capital stock corresponding to a predicted investment level at time t equals this predicted level divided by the depreciation rate of capital. We take 2% for the value of the depreciation rate of capital. This value corresponds to the depreciation rate of real estate as estimated by Harding et al. (2007) on AHS data.

Third, the dynamic panel data model does not estimate fixed effects, while these are important to make counterfactual predictions with a log-log specification. This is because an increase in the log. expenditure by 1 unit has a different impact on the finally predicted expenditure level depending on the initial value of the log. expenditure. To circumvent this problem, the micro-simulation is calibrated with the observed data. Hence, if the observed electricity (or gas) expenditure is 100 USD and the model prediction without climate change is 70 USD, then we take into account that there is an error term of 30 USD that needs to be accounted for before calculating a new expenditure level under climate change.

In addition, since the energy consumption model is dynamic, immediate shocks have resilient effects on energy consumption. We therefore need to estimate long term consumption. We use the standard econometric approach:

any change in the fitted values is multiplied by a factor of $1/(1-\hat{\gamma}_f)$, where $\hat{\gamma}_f$ is the estimated value for the coefficient of the lagged dependent variable.

Moreover, we are as much interested in the point estimates of the overall estimation as in the confidence intervals of the simulation. In particular, we want to know if there is some probability that adaptation costs are excessively high once we sum up investment and energy costs. This is why we compute point estimates and confidence intervals using a Monte-Carlo simulation: for each observation, we produce 200 draws of investments (separately in equipment and weatherisation). These draws are randomly obtained from the output of the investment models. We make random predictions based on the variance-covariance matrix associated with the vector of parameters estimated from the investment equations. We denote each equiprobable set of investments I_{it}^{Π} for individual i at time t and draw Π.

To derive observation-specific energy consumption estimates, we plug each simulated draw of investments into the electricity and gas consumption equations. Mathematically, we have $q_{ift}^{\Pi}=F_{if}^{\Pi}(I_{it}^{\Pi})$ where q_{ift}^{Π} is the resulting energy consumption level for individual i at time t and fuel f and draw Π. Importantly, q_{ift}^{Π} depends on I_{it}^{Π} but also on the functional form $F_{if}^{\Pi}()$ linking investments to the consumption of fuel f. This functional form is assumed to be random and specific to individual i and draw Π. Instead of using the set of parameters as estimated with the electricity and gas consumption specifications, we also take random draws of $F_f()$ based on the variance-covariance matrix of the parameters previously estimated with the energy consumption model.

In the end, we obtain 200 estimates per observation of electricity and gas consumption differentials between a scenario with no change in temperature and a scenario featuring a 1 Fahrenheit degree increase. Averaging these estimates provides us with point estimates. On the other hand, we also compute observation-specific confidence intervals. Their bounds are simply obtained by looking at the 2.5% lowest and highest draws. From the individual-specific confidence intervals, we can provide a conservative estimate of the sample confidence intervals by making the simplifying assumption that the bounds of the confidence intervals are reached for the same set of parameters across observations.[16]

5.2. Results

Table 6 first compares the average investment level of capital in equipment and weatherization between the temperatures as observed for our sample and the A2 scenario. It also provides the adjustments in electricity and gas expenditure, which take into account the predicted changed in capital stocks[17]. Finally, using a discount rate of 4% and a time-frame of 25 years, we provide an estimate of the total discounted cost of adaptation to climate change of the average housing unit.

[16] This assumption overestimates the average confidence interval for the sample but is necessary to ensure computable confidence intervals.

[17] We have disregarded any change in other indoor amenities and focus on energy-related investments.

Table 6: Estimated impact of the A2 scenario (2080-2099) on annual investments and energy expenditure for a representative US housing unit

Type of expenditures	Sample average 1985-2011	Variation under the A2 scenario		
		In level		In percent
		Mean	95% confidence interval	
Annual investment in equipment	\$147	+ \$121	[- \$50, + \$293]	+82%
• *For heating*		- \$113**	[- \$213, - \$14]	-
• *For cooling*		+ \$235***	[+ \$92, + \$377]	-
Annual investment in weatherization	\$417	- \$30	[- \$380, + \$320]	-7%
Annual electricity bill†	\$1,290	+ \$266**	[+ \$9; + \$525]	+21%
Annual gas bill†	\$998	- \$150***	[- \$288; - \$47]	-15%
Total discounted cost of adaptation†	-	+ \$3,947	[- \$4,773; + \$12,556]	-

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. 95% confidence intervals in brackets. †: for energy expenditure levels and net present values, main impacts are median impact, not mean impacts to avoid taking into account extreme values while using of a log-log specification: 1st and 99th percentile are explosive. Discounted cost of adaptation is calculated for 25 years with a 4% discount rate. All monetary numbers are in 2001 US \$.

We see that investments in equipment may increase drastically by 82% (+121 $). The reason is the massive need for air conditioning during very hot days (+235 $) which is only partially offset by a reduction in equipment for heating (-113 $). The level of weatherization seem to be less affected as we predict a moderate 30 $ decrease, which is not statistically significant.

The model predicts an increase in electricity bills by 21% and a decrease in gas expenditure by 15%. Although statistically significant at 5%, these estimates are characterized by large confidence intervals. The increase in electricity expenditures is estimated to lie between 1% and 41% and the reduction in gas expenditures between 5% and 29%.

The discounted cost of adaptation to climate change is estimated to be around $3,950, but this value is not statistically different from zero. To put this in context, this is equivalent to an annual expenditure of 210 $, which is approximatively 0.2% of the sample average annual household income or about 1.9% of the average purchase price of the housing units included in the sample (around 211,000 $). So this appears to be a small loss. If one uses as a conservative estimate the upper bound of the confidence interval, the cost would be $12,550. This is equivalent to an annual cost of about $750, equivalent to an increase in expenditure by around 0.85% of the average U.S. family budget, or the average gas expenditure in the sample.

Complementary simulation results are provided in Online Appendices B.7 and B.9. In Online Appendix B.7, we calculate the impacts of climate change on energy consumption instead of energy expenditure. We find that climate change leads to an increase in the electricity consumed by households (+16%), offset by a proportionally similar decrease in gas consumption (-16%). Because households consume roughly twice as much gas as electricity, the net effect is a non-statistically significant reduction in energy consumption by 7%.

In addition, a series of considerations need to be made about our final results. First, these cost estimates have been obtained from a moderate to high increase in land-based temperatures. Assuming smaller or wider changes in temperatures would lead to different estimates. Second, two phenomena are likely to lower adaptation costs. First, we have assessed the economic cost for households to adapt existing dwellings only. In 2013, there were around 990,000 building permits issued whereas the total stock of housing units was around 132 million[18]. After 25 years, the total number of new building permits should represent around 20% of the currently existing stock. For these new houses, adaptation to new temperatures will be considered at the development phase and, therefore, we could expect even smaller costs for new buildings than the ones estimated for existing units. This is likely to reduce the average cost of adaptation per household. Lower costs could also be fostered by technological improvements in heating and cooling systems, leading to either a reduction in investment costs or energy bills.

6. Conclusion

This research has developed a two-stage econometric approach to analyse adaptation of US homes to climate change. In the first stage, we have analysed the responsiveness of residential renovation efforts to climatic change. The results of our first stage have then been used in the second stage to predict how residential electricity and gas demand could evolve under climate change.

This research finds that total residential energy expenditures could slightly increase in the US as a result of climate change, with an estimated increase by 5.1%. However, residential energy demand would shift from gas to electricity with significant reallocation of energy use across the country: gas expenditure is expected to decrease sharply (-15%), electricity expenditure increase significantly (+21%).

But the main findings concern the cost of adaptation. We find that the cost for adapting to climate changes could only be 3,500 $ per household. This number

[18] USA Quick facts from the Census Bureau, consulted in March 2014: http://quickfacts.census.gov/qfd/states/00000.html

appears to be still moderately low when put in relation to U.S. average household income and the long-time span over which adaptation costs may incur.

Precautions should however be taken at interpreting our results, as we have assumed no economic growth or demographic evolution. Furthermore, we have assumed no change in the technologies available to households for space heating and air-conditioning, in terms of energy efficiency, but also in terms of fuel choice for space heating or air-conditioning: for example our results are conditional on gas not being used more often for air-conditioning. In addition, we only study the impact of climate change on existing homes whereas it will also influence the design of new homes. It could also accelerate the pace of renewal of the housing stock. Furthermore, we only considered the impact of temperatures and not the increased risk of flooding associated with sea level rise or any potential increase in liabilities due to a change in the frequency of disasters, for which large uncertainties remain.

Additionally, note that the US housing stock is relatively specific to the extent that gas consumption is high and air-conditioning is already present in many US homes: 86% of households in our sample declared having at least one air-conditioner at home, whereas 77% live in neighbourhoods connected to pipe gas. Thus, there is clearly a need in conducting similar studies on the relationship between climate change and home characteristics in other countries.

References

Alberini, Anna, Will Gans and Daniel Velez-Lopez, 2011. “Residential Consumption of Gas and Electricity in the U.S.: The Role of Prices and Income,” *Energy Economics*, Volume 33, Pages 870–881.

Auffhammer, M. and E. T. Mansur, 2014. "Measuring Climatic Impacts on Energy Expenditures: A Review of the Empirical Literature".

Auffhammer, Maximilian and Anin Aroonruengsawat. 2011. “Simulating the Impacts of Climate Change, Prices and Population on California’s Residential Electricity Consumption,” *Climatic Change*, Volume 109, Issue of first semester, Pages 191–210.

Auffhammer, Maximilian and Anin Aroonruengsawat. 2012 revision of: “Simulating the Impacts of Climate Change, Prices and Population on California’s Residential Electricity Consumption,” Energy Institute at Haas.

Barreca, Alan, Karen Clay, Olivier Deschênes, Michael Greenstone and Joseph S. Shapiro, 2016. “Adapting to Climate Change: The Remarkable Decline in the U.S. Temperature-Mortality Relationship Over the 20th Century?” Journal of Political Economy, 124(1): 105-159;

Blundell, Richard and Stephen Bond, 1998. “Initial conditions and moment restrictions in dynamic panel data models,” *Journal of Econometrics*, Volume 87, Issue 1, Pages 115–143

Blundell, Richard, Rachel Griffith and Frank Windmeijer, 2002. “Individual e-ffects and dynamics in count data models,” Journal of Econometrics, Volume 108, Pages 113-131.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2009. “Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates.” American Economic Review 99 (2): 198–204.

Dell, Melissa, Benjamin Jones and Benjamin Olken, 2009. "Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates," *American Economic Review*, Volume 99, Issue 2, Pages 198–204.

Dell, Melissa, Benjamin Jones and Benjamin Olken, 2012. "Temperature shocks and economic growth: Evidence from the last half century," *American Economic Journal: Macroeconomics*, Volume 4, Issue 3, Pages 66–95.

Dell, Melissa, Benjamin Jones and Benjamin Olken, 2014. “What Do We Learn from the Weather? The New Climate-Economy Literature,” *Journal of Economic Literature*, Volume 52, Issue 3, Pages 740–798.

Deschênes, Olivier and Michael Greenstone, 2011. “Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US,” *American Economic Journal: Applied Economics*, Volume 3, Pages 152–185.

Dixit, A. and R. Pindyck, *Investment under Uncertainty*, Princeton University Press, Princeton, N.J., 1994.

Doms, Mark and Timothy Dunne, 1998. “Capital Adjustment Patterns in Manufacturing Plants,“ Review of Economic Dynamics, Volume 1, Pages 409–429.

Gelain, Paolo and Kevin J. Lansing (2014) “House prices, expectations, and time-varying fundamentals,” *Journal of Empirical Finance*, Volume 29, Pages 3–25.

Harding, John P., Stuart S. Rosenthal and C.F. Sirmans, 2007. "Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model," *Journal of Urban Economics*, Volume 61, Issue 2, Pages 193–217.

Helms, Andrew C., 2003. “Understanding gentrification: an empirical analysis of the determinants of urban housing renovation,” *Journal of Urban Economics*, Volume 54, Pages 474–498.

Honoré, Bo E., 1992. ”Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects,“ *Econometrica*, Volume 60, Issue 3, Pages 533–565.

Intergovernmental Panel on Climate Change (IPCC), 2001. *Climate change 2001: impacts, adaptation and vulnerability*. IPCC Working Group II.

Intergovernmental Panel on Climate Change (IPCC), 2013. *Fifth Assessment Report: Climate Change 2013* (AR5).

Mansur, E.T., R. Mendelsohn and W Morrison, 2008. “Climate Change Adaptation: A Study of Fuel Choice and Consumption in the US Energy Sector,” *Journal of Environmental Economics and Management*, Volume 55, Issue 2, 175-193.

National Association of Home Builders, 2010. “Breaking Down House Price and Construction Costs”

World Bank. 2010. The Cost to Emerging countries of Adapting to Climate Change: New Methods and Estimates. Washington, DC: World Bank.

Main appendices

A.1: Contemporaneous and lagged values of climate variables

Instead of using the expected values for heating and cooling degree days and precipitations, we use contemporaneous values. Results fail to capture the impact of cooling degree days on equipment, probably because only repeated heat waves are necessary to increase investments in air conditioning. Likewise, the model fails to capture the impact of heating degree days on insulation.

Table7: Linear investment models using contemporaneous values

Type of investment	Equipment	Weatherization
Heating degree days	0.0683**	0.0464
	(2.25)	(0.69)
Cooling degree days	0.0637	0.147
	(1.23)	(1.40)
Precipitations	0.000311	-0.00301
	(0.08)	(-0.36)
No. people in unit	-3.393	35.77
	(-0.34)	(1.52)
Connection to pipe gas	97.60*	137.4
	(1.73)	(1.46)
Observations	47,554	45,366

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported.

Another possibility consists in using a distributed lag model, where the model includes lagged values for heating and cooling degree days. This type of model confirms that lagged values need to be taken into account. Effects are close to the ones of the main model.

Table8: Distributed lag model for investments

Type of investment	Equipment	Weatherization
Heating degree days:	0.0435	0.00220
	(1.21)	(0.03)
First year lag	0.0371	0.153**
	(1.10)	(2.10)
Second year lag	-0.0374	-0.00179
	(-0.95)	(-0.02)
Third year lag	0.0407	0.0794
	(1.09)	(0.95)
Cooling degree days:	0.0905	0.199*
	(1.56)	(1.68)
First year lag	0.0354	-0.132
	(0.69)	(-1.11)
Second year lag	0.138**	0.121
	(2.39)	(1.01)
Third year lag	-0.00692	0.0484
	(-0.14)	(0.48)
Precipitations:	0.000224	0.00658
	(0.05)	(0.71)
First year lag	-0.00530	0.000928
	(-1.20)	(0.09)
Second year lag	0.00380	0.0248***
	(0.93)	(2.79)
Third year lag	-0.000681	0.00741
	(-0.15)	(0.77)
No. people in unit	-7.496	39.67*
	(-0.74)	(1.67)
Connection to pipe gas	97.24*	148.6
	(1.71)	(1.57)
4-year cumulative impact:		
Heating degree days	0.0839	0.2328*
	(1.48)	(1.91)
Cooling degree days	0.2574**	0.2360
	(2.18)	(1.00)
Precipitations	-0.0020	0.0400*
	(-0.23)	(1.91)
Observations	46,505	44,353

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies.

A.3: Using temperature bins, by US Region

For both hot and cold regions, we run the linear investment models with the expected amount of days falling within temperature and precipitation bins using a value of lambda equal to around 0.31.

Table 9: Fixed-effect investment models using expected days falling in each temperature bin, with separate effects for hot and cold regions

Type of investment	Equipment		Weatherization	
U.S. Regions	Cold	Hot	Cold	Hot
Expected days with temperature:				
Below 10°F	-5.923	-77.14^{**}	-7.871	41.42
	(-0.61)	(-2.27)	(-0.37)	(0.54)
Between 10-20°F	8.174	4.220	-6.174	35.41
	(0.93)	(0.24)	(-0.33)	(0.98)
Between 20-30°F	-6.015	9.098	-25.53^{*}	49.95^{**}
	(-0.82)	(1.02)	(-1.73)	(2.56)
Between 30-40°F	1.699	14.93^{***}	-18.70	36.88^{***}
	(0.27)	(2.60)	(-1.31)	(3.23)
Between 40-50°F	4.577	-0.415	-14.75	6.597
	(0.70)	(-0.11)	(-1.11)	(0.78)
Between 50-60°F	-2.111	3.775	-4.433	10.42^{*}
	(-0.38)	(1.37)	(-0.39)	(1.71)
Between 60-70°F	-	-	-	-
Between 70-80°F	0.393	0.201	-12.61	3.319
	(0.09)	(0.08)	(-1.33)	(0.62)
Between 80-90°F	5.031	5.837^{*}	-30.47^{**}	6.896
	(0.66)	(1.73)	(-2.13)	(1.01)
Above 90°F	-135.5	17.07^{**}	-67.85	20.95^{*}
	(-0.96)	(2.34)	(-0.21)	(1.78)
Expected days with precipitations:				
No precipitation	-0.00739	0.680	0.725	1.956
	(-0.01)	(0.46)	(0.23)	(0.62)
Between 0-50mm	-2.354	-5.016	-6.397	-10.61
	(-0.51)	(-0.90)	(-0.68)	(-1.04)
Between 50-100mm	0.517	-0.621	9.598	9.254
	(0.10)	(-0.10)	(0.81)	(0.80)
Between 100-200mm	-7.596	4.794	-6.792	-1.775
	(-1.13)	(0.75)	(-0.39)	(-0.13)
Above 200mm	-8.425	0.286	62.17^{*}	24.66
	(-0.57)	(0.02)	(1.85)	(0.74)
No. people in unit	179.9^{**}	22.15	16.78	261.8^{*}
	(2.19)	(0.29)	(0.15)	(1.81)
Connection to pipe gas	-5.923	-77.14^{**}	-7.871	41.42
	(-0.61)	(-2.27)	(-0.37)	(0.54)

Observations	24,173	23,561	22,607	22,911

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies.

A.4: Expenditure models with bins

We report below the results when temperature and precipitation bins are used in the energy expenditure model. For concision, only the coefficients for the bins are reported since all the other coefficients remained stable. We report first the results for the entire sample, and then separately for hot and cold regions.

Table 10: Blundell-Bond estimation of energy expenditure models with bins

Type of fuel	Electricity	Gas
Expected days with temperature:		
Below 10°F	0.0000433	-0.00208**
	(0.05)	(-2.18)
Between 10-20°F	0.00199**	0.00522***
	(1.97)	(4.98)
Between 20-30°F	-	0.00386***
		(4.63)
Between 30-40°F	0.00127*	0.00348***
	(1.89)	(3.72)
Between 40-50°F	0.00252***	0.00230***
	(4.55)	(4.90)
Between 50-60°F	0.000635	0.000108
	(1.27)	(0.35)
Between 60-70°F	0.00142***	0.000466
	(2.77)	(1.49)
Between 70-80°F	0.00130***	0.00102***
	(2.67)	(2.85)
Between 80-90°F	0.00216***	-
	(3.66)	
Above 90°F	0.00261***	0.00287***
	(3.47)	(3.23)
Expected days with precipitations:		
No precipitation		
Between 0-50mm	0.000170	-0.000213
	(1.35)	(-1.34)
Between 50-100mm	0.000274	0.00159*
	(0.57)	(1.86)
Between 100-200mm	0.000139	0.00474***
	(0.21)	(4.86)
Above 200mm	0.00222***	0.00122
	(2.76)	(1.40)

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include all the variables used in the base specification. Constant terms are not reported. Lagged dependent variable is instrumented using first lag of average commuting time and interaction between commuting time and square footage of unit at time of purchase. All (dependent, independent and instrumental) variables are in logarithm.

Table 11: Blundell-Bond estimation of energy expenditure models with bins, by region

Type of fuel	Electricity		Gas	
U.S. Regions	Cold	Hot	Cold	Hot
Expected days with temperature:				
Below 10°F	-0.00162	0.00724	-0.000485	-0.00893*
	(-1.47)	(1.20)	(-0.35)	(-1.74)
Between 10-20°F	0.000385	0.00772**	0.00538***	-0.00483
	(0.32)	(2.00)	(3.54)	(-1.16)
Between 20-30°F	-	-	0.00447***	0.00290
			(3.47)	(1.59)
Between 30-40°F	0.000810	0.00480***	0.00317**	0.00566***
	(0.94)	(3.03)	(2.10)	(4.08)
Between 40-50°F	0.00154**	0.00482***	0.00422***	0.00226***
	(2.01)	(3.59)	(3.37)	(3.80)
Between 50-60°F	0.00179**	0.00402***	0.00264*	0.000601
	(1.96)	(3.12)	(1.94)	(1.46)
Between 60-70°F	0.00202**	0.00390***	0.00238*	0.0000716
	(2.32)	(3.13)	(1.86)	(0.17)
Between 70-80°F	0.00189**	0.00396***	0.00304**	0.000652
	(2.46)	(3.17)	(2.27)	(1.49)
Between 80-90°F	0.00153*	0.00539***	-	-
	(1.86)	(3.69)		
Above 90°F	0.0533	0.00644***	0.0417	0.00298***
	(0.97)	(4.08)	(0.38)	(2.92)
Expected days with precipitations:				
No precipitation	-	-	-	-
Between 0-50mm	-0.000658***	0.000545**	-0.000909***	0.000115
	(-3.68)	(2.46)	(-3.53)	(0.44)
Between 50-100mm	-0.000999	-0.000625	0.000331	-0.000632
	(-1.35)	(-0.76)	(0.38)	(-0.44)
Between 100-200mm	-0.00284***	0.00343***	0.00391***	0.00614***
	(-3.12)	(3.74)	(3.44)	(3.96)
Above 200mm	-0.00279**	0.00237**	-0.00244	0.00421***
	(-2.42)	(2.27)	(-1.59)	(3.15)

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. Lagged dependent variable is instrumented using first lag of average commuting time and interaction between commuting time and square footage of unit at time of purchase. All (dependent, independent and instrumental) variables are in logarithm.

A.5: Method used to recover the capital amounts

Let's note $k_{i,t}$ the amount of capital embodied in the home of household i at time t, net of any observed home improvement $I_{h,i,t}$. For the years in which the housing units are either sold or built, $k_{i,t}$ is known to us. For the years following the sale/construction of the house, we input the value of $k_{i,t}$ by applying a depreciation rate on housing capital:

$$k_{i,t} \approx (1-\delta)^{\tau_i} k_{i,t-\tau_i}$$

τ_i represents the observed date of construction or sale. We take 2% for the value of the depreciation rate of past investments (i.e. $\delta = 2\%$). This value corresponds to the depreciation rate of real estate as estimated by Harding *et al.* (2007) on AHS data. Additionally, for the years that precede a sale and for which we have no previous information on the initial capitalized investments net of home improvements, we infer it from the sales price of the home at a later date:

$$k_{i,\tau_i-s} \approx \frac{k_{i,\tau_i}}{(1-\delta)^s}$$

s represents the lag between the observed purchase and the time of interest for the calculation of k. This technique gives us a proxy of the amount of capital in a home before home investments are made provided that we observe at least one sale or the construction cost of the unit.

We expect our approximation of $k_{i,t}$ to be representative of the value of all the services delivered by the housing unit net of any observed investment $I_{h,i,t}$. We however want to distinguish the initial stock of capital associated with equipment and the initial stock of capital associated with weatherization from the rest. To do so, we use the information provided by the National Association of Home Builders (NAHB, 2010) on construction costs. According to the NAHB, 20.3% of the construction cost of a single-family unit is due to the lot price. Furthermore, the NAHB also provides a breakdown of the construction cost of a home according to the part of the unit that is considered. In particular, heating, ventilation and air-conditioning systems represent 4.0% of the construction cost, and appliances 1.6% in average. We therefore proxy the initial stock of capital in equipment, net of any home improvement after τ_i, by evaluating the share of $k_{i,t}$ that is most likely to have been allocated to equipment at the time of construction based on NAHB (2010). This leads us to apply the following formula:

$$k_{1,i,t} = k_{i,t} * (1 - 20.3\%) * (4\% + 1.6\%)$$

In the equation above, $k_{1,i,t}$ is the capitalized investments in type 1 (equipment) at time t for household i, net of any improvement performed to the home after τ_i. We can likewise assess the capitalized investments net of home improvements for insulation, storm doors and windows, and all the other home improvements covered with our data.[19]

Once the capitalized investments net of home improvements after τ_i have been calculated for all the three types of homes improvements, we add the value of all the home improvements performed in the house since the last purchase (τ_i) or withdraw the sum of all the home improvements done between time t and the observed future purchase (in τ_i) to proxy the value of capitalized investments in a specific type h of housing services at time t:

$$K_{h,i,t} = \begin{cases} k_{h,i,t} + \sum_{s=\tau_i}^{t} I_{h,i,s}(1-\delta)^{t-s} \; if \; t \geq \tau_i \\ k_{h,i,t} - \sum_{s=t}^{\tau_i} I_{h,i,s}(1-\delta)^{\tau_i - s} \; if \; t < \tau_i \end{cases}$$

[19] According to NAHB (2010), Insulation is 1.5% of construction costs, windows represent 2.8%, exterior doors 0.9%, framing and trusses 15.6%, roof shingles 3.8% and siding 5.8%. To assess the initial capital for all the other homes improvements, we consider that it correspond to the remaining share, excluding outdoor features and fees, i.e. lanscaping and sodding (3.2%), wood decks or patios (0.9%), asphalt driveways (1.4%), building fees (1.9%) and impact fees (1.4%).

To get comparable values of $K_{h,i,t}$ through time, $k_{h,i,t}$ and $I_{h,i,s}$ are deflated using the US consumers' price index (CPI) of the Bureau of Labor Statistics.

Online supplementary appendices

B.1: Left-censored investment models

We are using a panel data tobit model instead of a fixed effect linear model to estimate investments. This approach is relevant because investments are only observed with a positive value. Hence, we complement our linear model with a latent variable approach which is similar to that of Helms (2003), except that we fully take advantage of the panel structure of the data.[20] Mathematically, we assume that investment I_{iht} depends on a latent variable I_{iht}^{*} which is defined by:

$$I_{iht} = \begin{cases} I_{iht}^{*} \; if \; I_{iht}^{*} > 0 \\ 0, \text{otherwise} \end{cases}$$

We then assume that I_{iht}^{*} is equal to the right hand-side of Eq. (1). We estimate this model using the estimator developed by Honoré (1992) for panel data tobit models. This estimator accounts for household-specific fixed effects (u_i) and therefore fully exploits the panel structure of our data.

Table 17 provides the results of the investment equations when panel tobit models are applied. The absolute value of the coefficients between a linear model and a panel tobit model are different due to the change in model used. However, significance levels are similar between the two models (i.e. Table 17 versus Table 3), along with the relative impact of heating and cooling degree days on investments.

[20] An alternative approach would consist in using two stages instead of a latent variable model. The first stage would be a logit or a probit model predicting the probability of investment, and the second stage a linear equation for the amount invested, conditional of an investment being made. We tried to adopt this approach but could not obtain satisfying results with it, principally because we do not have many observations with a positive investment. It follows that we rarely observe two home improvements of the same type for a given household. In the second stage, this approach implies using a fairly restricted sample of observations, implying significant losses in terms of efficiency.

Table 12: Results of investment equations using a panel tobit model

Type of investment	Equipment	Weatherization
Expected heating degree days	1.769**	1.806**
	(2.04)	(2.22)
Expected cooling degree days	3.570**	1.674
	(2.39)	(1.12)
Expected precipitations	-0.0353	0.0685
	(-0.30)	(0.54)
No. people in unit	-48.72	227.4*
	(-0.41)	(1.70)
Connection to pipe gas	1120.9	978.1
	(1.54)	(1.35)
Observations	53,393	51,396

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported.

This model is not used as our base specification because it does not produce estimates of fixed effects that are necessary to make accurate predictions at the simulation stage.

B.2: Using observations that benefitted from a public grant or a loan

Investments in space heating, air cooling or weatherization are also influenced by energy efficiency policies. If these policies are correlated with climate shocks or expectations, our estimates of the impact of climate adaptations on home improvements could be biased. In the base specification, we exclude the observations that have benefitted from public support, considering that only around 2% of home improvements actually benefitted government grand or loans in our sample.

Alternatively, this piece of information can be used as a dummy variable in the home improvement models. However, this additional variable is likely to be endogenous. The reason is that it measures with error the availability of policies promoting energy efficiency at local level.[21] We report here alternative specifications where we use a control function approach to treat this endogeneity. We use as an instrument the share of households, within the same metropolitan statistical area and year, who benefitted from government support to perform investments in other amenities. This captures the likelihood of access to government support at local level at time *t* for investment *h*. This factor is however exogenous to specific households since it does not depends on household's *i* characteristics. Standard identification tests using a linear, 2SLS model, have been performed and corroborate the validity of this instrument (they are not reported for the sake of concision). Only little differences in results are found with this specification and our base specification.

[21] In particular, this variable only provides information for the households that actually performed alterations. For the other households, we do not know if they could have had access to government support or not. In addition, this is a binary variable whereas household choices are driven by the size of subsidies.

Table 13: Fixed-effect investments models using observations that benefitted from a public grant or a loan

Type of investment	Equipment	Weatherization
Benefitted from a public grant or a loan	-1050.7	-2406.5
	(-1.54)	(-1.46)
Expected heating degree days	0.165**	0.415***
	(2.30)	(2.62)
Expected cooling degree days	0.333**	0.271
	(2.55)	(1.03)
Expected precipitations	-0.00135	0.0145
	(-0.13)	(0.59)
No. people in unit	-4.897	29.47
	(-0.50)	(1.26)
Connection to pipe gas	63.00	108.5
	(1.09)	(1.15)
Weak identification test:		
Kleibergen-Paap rk Wald F statistic	34	28
Overidentification test†:		
Hansen J statistic (p-value)	0.50	0.53
Observations	47,434	45,214

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. †: Over-identification tests are run using the squared value of the instrument as a second instrument in a separate regression.

B.3: Fixed effect investment models with lagged capital stocks

In the base model, we have excluded previous capital already invested in the categories of investment. In this specification, we account for the impact of past capital in category *h* for investments in category *h*.

This creates an econometric difficulty: fixed effect models rely on the strict exogeneity of independent variables but the stock of capital is predetermined by past investments. To solve the problem, we use the discounted value of the house since its purchase, excluding any new investment made by the household, as instrumental variable in a 2SLS setting. This instrument is relevant since new investments will be correlated with the depreciation of the initial stock of capital in the house at the moment of purchase. Strict exogeneity of the instrument requires that it is uncorrelated with any shock on the investments performed by household i after purchase. This is necessarily the case for this instrument because the depreciation of the capital stock made before the purchase cannot depend on the time-varying unobserved characteristics of households after purchase. Overall results lose precision as compared to the base model, probably because of the use of 2SLS but coefficients remain in line with the ones obtained in the main model.

Table 14: Results of 2SLS investment models with lagged capital stocks

Type of investment	Equipment	Weatherization
Expected heating degree days	0.173	0.427
	(1.39)	(1.50)
Expected cooling degree days	0.365	0.367
	(1.61)	(0.80)
Expected precipitations	0.00133	-0.0398
	(0.07)	(-0.93)
No. people in unit	-14.96	-25.48
	(-0.89)	(-0.64)
Connection to pipe gas	40.03	153.8
	(0.51)	(1.06)
Capital in equipment at t-1	-0.0664	
	(-1.03)	
Capital in weatherization at t-1		-0.0335
		(-1.24)
Weak identification test: Kleibergen-Paap rk Wald F statistic	166	588
Overidentification test†: Hansen J statistic (p-value)	0.28	0.87
Observations	17,018	16,150

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. †: Over-identification tests are run using the squared value of the instrument as a second instrument in a separate regression.

B.4: Investment model for other indoor amenities

We have run a fixed-effect model to predict the amounts invested in other indoor amenities.

Table 15: Fixed-effect investment model for other indoor amenities

Type of investment	Other indoor amenities
Expected heating degree days	-0.341
	(-0.93)
Expected cooling degree days	0.0683
	(0.11)
Expected precipitations	0.106^{*}
	(1.83)
No. people in unit	97.06^{*}
	(1.82)
Connection to pipe gas	108.4
	(0.51)
Observations	50,334

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported.

B.5: Arellano-Bond estimator of energy expenditure model

The table below displays the results of the dynamic panel data model of energy expenditure using an Arellano-Bond estimator. The model fails to significantly capture the impact of the lagged dependent variable and brings imprecise results.

Table 16: Arellano-Bond estimation of energy expenditure models

Type of fuel	Electricity	Gas
Lagged dependent variable	0.0252	0.135
	(0.14)	(0.58)
Capital in equipment	-0.0572	0.0284
	(-1.18)	(0.42)
x heating fuel is electricity	0.0104***	
	(3.79)	
x AC fuel is electricity	0.00493**	
	(2.11)	
x heating fuel is gas		0.0217***
		(5.40)
x AC fuel is gas		0.00314
		(0.81)
Capital in weatherization	0.0771	-0.0738
	(0.86)	(-0.65)
Capital in other amenities	0.0350	0.0380
	(0.55)	(0.57)
Heating degree days	-0.0230	0.0778
	(-0.66)	(1.12)
Cooling degree days	0.0241	0.0636**
	(0.83)	(2.09)
Precipitations	0.0228	0.00990
	(1.32)	(0.46)
Connection to pipe gas	-0.0210	0.129
	(-0.68)	(1.52)
No. people in unit	0.190***	0.0704***
	(8.20)	(2.88)
Average commuting time	0.0153*	0.00555
	(1.95)	(0.52)
Observations	11192	7688
Hansen test	0.38	0.57

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. Capital levels are instrumented using their lagged values. Lagged dependent variable is instrumented using first lag of average commuting time. Because of multicollinearity, interaction between commuting time and square footage of unit at time of purchase was dropped. All (dependent, independent and instrumental) variables are in logarithm.

B.6: Fixed-effect energy expenditure model with AR(1) disturbance

We use fixed effect linear models with an AR(1) disturbance to test if the bias found in the fixed effect specification can be imputed to omitting the lagged dependant variable in the model. Such models are known to be in-between dynamic panel data models and a fixed effect models, but do not tolerate endogenous variables. We simply run models which exclude capital levels and find expected correlations between temperatures and energy expenditure.

In fact these estimates look like those obtained of a fixed effect specification, suggesting weak instrumentation (see Appendix 11).

Table 17: Fixed-effect energy expenditure models with AR(1) disturbance term

	Equipment	Weatherization
Heating degree days	0.443***	0.541***
	(29.31)	(26.47)
Cooling degree days	0.206***	0.188***
	(18.15)	(12.89)
Precipitations	0.131***	0.0614***
	(12.60)	(4.40)
No. people in unit	0.183***	0.0803***
	(16.41)	(5.38)
Connection to pipe gas	0.000576	0.193***
	(0.03)	(4.72)
Observations	35,933	26,134

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies but capital variables have been excluded. Constant terms are not reported.

B.7: Using energy consumption instead of energy expenditure in the energy models

Instead of using expenditure levels, we run a model on electricity and gas consumption levels. We obtain consumption levels by dividing energy expenditures by the average price of fuels in each U.S. State. The energy price data is taken from the State Energy Data System administered by the US Energy Information Administration. The data includes information on residential and industrial energy prices for each US State from 1985 to 2011. We combine the energy price data with the AHS data by matching the metropolitan statistical areas of the AHS with the State-level information of the SEDS. Each time at MSA is situated on more than one State, average price values are obtained by calculating the average energy price corresponding to the different States on which a metropolitan statistical area is overlapping.

In these alternative models where energy consumption is used as a dependent variable, we include energy prices as control variables. These are endogenous variables due to the simultaneous determination of energy prices and residential energy demand. To deal with such endogeneity, we construct instruments based on pre-sample information on energy prices between 1970 and 1983. The use of pre-sample information to construct instruments is a strategy that has been implemented recently in several studies, in particular in Blundell, Griffith and Windmeijer (2002). More precisely, we predict State-level gas and electricity prices for 1985-2011 with the pre-sample information we have. For each metropolitan statistical area, we run two Autoregressive Integrated Moving Average (ARIMA) models on the energy data, using only the years before 1983: one for residential electricity prices and one for residential gas prices. We use a model with three autoregressive orders for residential electricity and gas prices. They correspond to the best fits of any ARIMA model with our data. In addition, we include four time dummies for 1973, 1974, 1979 and 1980. These dates correspond to the first and second petroleum shocks for which we needed to control. We therefore have:

$$p_{rft} = a_{0,r} + a_{1,r}p_{rft-1} + a_{2,r}p_{rft-2} + a_{3,r}p_{rft-3} + \sum_{x} a_{x,r}(t = x) + e_{rft}$$

With p_{rft} is the price of fuel f in area r at time t. x corresponds to the dates of the petroleum shocks (1973, 1974, 1979 and 1980). $a_{0,r}$, $a_{1,r}$, $a_{2,r}$, $a_{3,r}$ and $a_{x,r}$ are area-specific parameters estimated by the ARIMA models. e_{rft} is an error term.

We take the predictions of these models for 1985-2011 and use them as instruments. By construction, they correspond to a probable trend in energy prices based on pre-sample information. In other words, we extrapolate trends in electricity and gas prices for 1985-2011 based on the information available in 1983. These instruments will certainly be correlated with observed prices from 1985 to 2011: the prediction for 1985 will be correlated with the observed price for 1985, the prediction for 1987 with the observed price for 1987 and so on. They however only correspond to a trend based on prior information: they are unrelated with any shock on energy demand for 1985-2011.

Results using energy consumption are similar to the ones obtained with energy expenditure.

Table 18: Blundell-Bond estimation of energy consumption models, controlling for energy prices

Type of fuel	Electricity consumption	Gas consumption
Lagged dependent variable	0.568***	0.340**
	(5.90)	(2.34)
	-0.415***	0.312***
	(-3.83)	(4.28)
	0.0954*	-0.386***
	(1.76)	(-3.66)
Capital in equipment	-0.0108	0.00217
	(-0.81)	(0.12)
x heating fuel is electricity	0.00824***	
	(5.30)	
x AC fuel is electricity	0.00702***	
	(3.85)	
x heating fuel is gas		0.0354***
		(5.44)
x AC fuel is gas		0.00289
		(1.12)
Capital in weatherization	0.0263	-0.0647*
	(1.16)	(-1.93)
Capital in other amenities	0.0254	0.0663**
	(0.98)	(1.98)
Heating degree days	0.0177**	0.291***
	(2.12)	(4.74)
Cooling degree days	0.0681***	0.0789***
	(3.31)	(4.93)
Precipitations	0.0432***	0.0865***
	(3.60)	(3.20)
Connection to pipe gas	-0.0503***	0.101*
	(-2.87)	(1.74)
No. people in unit	0.131***	0.0977***
	(6.18)	(5.17)
Average commuting time	-0.0538***	-0.0929***
	(-3.35)	(-3.76)
x sq. footage at time of Purchase	0.0586***	0.0878***
	(4.21)	(3.67)
Observations	19,598	13,718
Hansen test	0.32	0.31

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. Lagged dependent variable is instrumented using first lag of average commuting time and interaction between commuting time and square footage of unit at time of purchase. We also instrument energy prices using predictions for this prices based on pre-sample data. All (dependent, independent and instrumental) variables are in logarithm.

Similar regressions have been run separately for hot and cold regions but are not reported for the sake of concision. We have furthermore run the simulation on the impacts of climate change on these energy consumption models to estimate the impacts in terms of energy consumptions. Results are provided in the table below.

Table 19: Estimated impact of global warming on energy consumption

Impact of global warming (A2 scenario, 2080-2099)	All sample	Cold regions	Hot regions
Relative change in electricity consumption	+16%* (-2% ; +34%)	+25%** (+5% ; +45%)	+20%* (-4%; +45%)
Relative change in gas consumption	-16%** (-3% ; -33%)	-27%** (-57% ; -6%)	-16%** (-1%; -31%)

* p < 0.1, ** p < 0.05, *** p < 0.01. 95% Confidence intervals in brackets. Main impacts are median impact, not mean impacts to avoid taking into account extreme values while using a log-log specification.

B.8: Energy expenditure model with no capital variable

Table 20: Blundell-Bond estimation of energy expenditure models

Type of fuel	Electricity		Gas	
Lagged dependent variable	0.558***	0.571***	0.359**	0.262*
	(5.35)	(5.75)	(2.35)	(1.67)
Heating degree days	-0.00201	0.00231	0.257***	0.266***
	(-0.34)	(0.39)	(4.29)	(4.67)
Cooling degree days	0.0876***	0.0646***	0.0475***	0.0534***
	(3.13)	(2.87)	(3.47)	(3.83)

Table 21: Blundell-Bond estimation of energy expenditure model without capital stocks

Type of fuel	Electricity	Gas
Lagged dependent variable	0.449***	0.235
	(4.86)	(1.64)
Heating degree days	-0.0155***	0.284***
	(-2.85)	(5.36)
Cooling degree days	0.0902***	0.0296***
	(4.38)	(2.88)
Precipitations	0.0594***	0.100***
	(5.12)	(4.71)
Connection to pipe gas	-0.129***	0.132**
	(-5.70)	(2.41)
No. people in unit	0.152***	0.109***
	(7.52)	(5.72)
Average commuting time	-0.100***	-0.120***
	(-4.93)	(-4.58)
x sq. footage at time of Purchase	0.104***	0.125***
	(5.46)	(4.68)
Observations	22,494	15,805
Hansen test	0.32	0.23

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. Lagged dependent variable is instrumented using first lag of average commuting distance and interaction between commuting distance and square footage of unit at time of purchase. All (dependent, independent and instrumental) variables are in logarithm.

Similar regressions have been run separately for hot and cold regions but are not reported for the sake of concision. We have furthermore run the simulation on the impacts of climate change with these energy models with no capital variable. Results are provided in Table 11.

B.9: Bundling all capital variables in one category

Instead of separating different types of capital, it is possible to bundle all capital variables into one category. Below, we present the results for both the investment and energy expenditure models in this case. The capital in the house at time t has been instrumented with its lagged value in the energy expenditure model.

Table 22: Fixed-effect investment model assuming one time of capital

Type of investment	Any investment in the house
Expected heating degree days	0.115
	(0.26)
Expected cooling degree days	0.503
	(0.67)
Expected precipitations	0.129*
	(1.80)
No. people in unit	135.7**
	(1.99)
Connection to pipe gas	376.6
	(1.40)
Observations	48,035

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported.

Table 23: Blundell-Bond estimation of energy expenditure model with one uniform capital stock

Type of fuel	Electricity	Gas
Lagged dependent variable	0.565***	0.437**
	(5.43)	(2.51)
Capital in the house	0.0515***	0.0537***
	(3.11)	(3.02)
Heating degree days	-0.00201	0.226***
	(-0.34)	(3.33)
Cooling degree days	0.0861***	0.0433***
	(3.08)	(2.92)
Precipitations	0.0547***	0.0778***
	(3.56)	(2.73)
Connection to pipe gas	-0.103***	0.149***
	(-4.09)	(2.68)
No. people in unit	0.129***	0.0827***
	(5.85)	(3.82)
Average commuting time	-0.0602***	-0.0710***
	(-3.21)	(-2.78)
x sq. footage at time of Purchase	0.0634***	0.0707***
	(3.90)	(2.65)
Observations	19,598	13,718
Hansen test	0.24	0.14

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models include household fixed effects and time-dummies. Constant terms are not reported. Lagged dependent variable is instrumented using first lag of average commuting distance and interaction between commuting distance and square footage of unit at time of purchase. All (dependent, independent and instrumental) variables are in logarithm.