G Model
RESEN-1042; No. of Pages 22
ARTICLE IN PRESS

Resource and Energy Economics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Resource and Energy Economics

journal homepage: www.elsevier.com/locate/ree

Chinese residential electricity consumption: Estimation and forecast using micro-data☆

Jing Cao[a,b], Mun Sing Ho[b,d], Yating Li[c,f], Richard G. Newell[d,e,*], William A. Pizer[c,e,f]

[a] School of Economics and Management, Tsinghua University, China
[b] Harvard China Project on Energy, Economy and the Environment, Harvard University, USA
[c] Energy Initiative, Duke University, United States
[d] Resources for the Future, 1616 P St NW, Washington DC, 20036, United States
[e] National Bureau of Economic Research, United States
[f] Sanford School of Public Policy, Duke University, United States

ARTICLE INFO

Article history:
Received 2 December 2016
Received in revised form 7 June 2017
Accepted 5 October 2017
Available online xxx

JEL classification numbers:
Q41
Q47
O44

Keywords:
Electricity
Demand
Estimation
Forecast
Residential
China

ABSTRACT

Based on econometric estimation using data from the Chinese Urban Household Survey, we develop a preferred forecast range of 85–143 percent growth in residential per capita electricity demand over 2009–2025. Our analysis suggests that per capita income growth drives a 43% increase, with the remainder due to an unexplained time trend. Roughly one-third of the income-driven demand comes from increases in the stock of specific major appliances, particularly AC units. The other two-thirds comes from non-specific sources of income-driven growth and is based on an estimated income elasticity that falls from 0.28 to 0.11 as income rises. While the stock of refrigerators is not projected to increase, we find that they contribute nearly 20 percent of household electricity demand. Alternative plausible time trend assumptions are responsible for the wide range of 85–143 percent. Meanwhile we estimate a price elasticity of demand of −0.7. These estimates point to carbon pricing and appliance efficiency policies that could substantially reduce demand.

1. Introduction

China's energy demand has grown rapidly in the past decade. Despite slower recent energy consumption growth as economic growth decelerates, a significant share of rising global consumption remains concentrated in China, among the other Asian countries. In 2014, China (+2.6%) and India (+7.1%) recorded the largest national increments to global energy

☆ This paper is supported by the National Science Foundation of China (project codes: 71422013 and 71173130), the Harvard Global Institute, the Hang Lung Center for Real Estate of Tsinghua University, in part by Volvo Group in a research project of the Research Center for Green Economy and Sustainable Development, Tsinghua University, the Tsinghua University Initiative Scientific Research Program and the Energy Initiative, Duke University.
* Corresponding author at: Resources for the Future, 1616 P St NW, Washington DC, 20036, United States.
E-mail address: newell@rff.org (R.G. Newell).

https://doi.org/10.1016/j.reseneeco.2017.10.003
0928-7655/

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

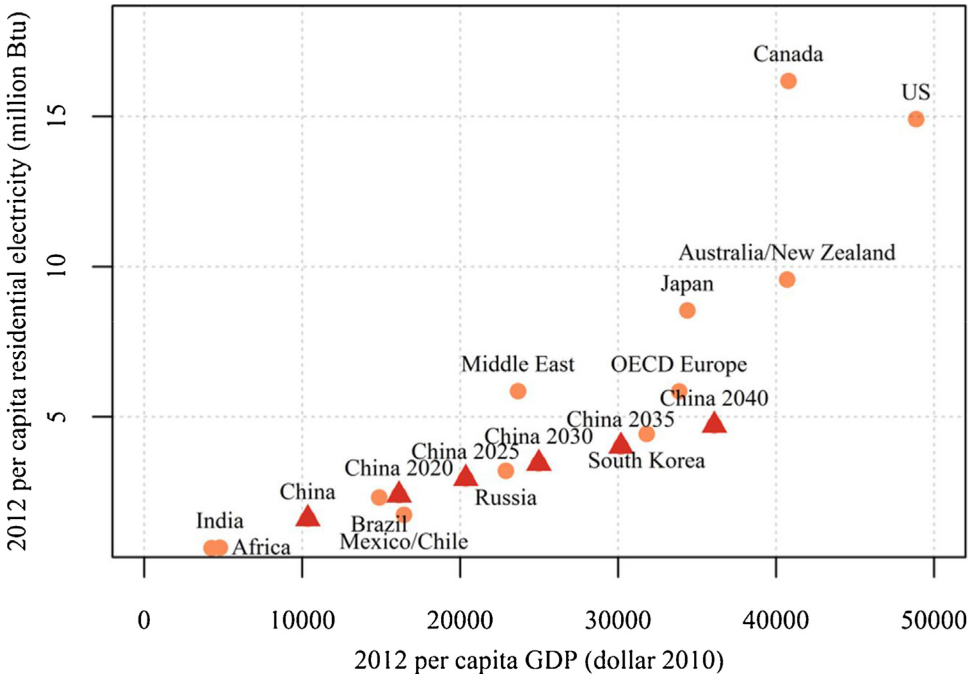

Fig. 1. Residential electricity per capita versus per capita GDP in 2012, alongside projections for China.
Source: EIA, International Energy Outlook 2016

consumption.[1] As a result of the high energy consumption and relatively heavy dependence on coal (nearly 66% of total primary energy, EIA), China overtook the United States as the world's leading emitter of greenhouse gases in 2006. The future growth path of energy consumption is relevant not only for projecting world energy market outcomes, but also for the strategies to tackle both local air pollution and global climate change.

This paper focuses on modeling and forecasting Chinese residential electricity demand using household survey data to 2025. Many previous forecasts of China's electricity demand have focused on projecting the total consumption using aggregate economic growth trends. Among sectors, the industrial sector, accounting for more than two-thirds of China's electricity consumption, attracts the most attention. In comparison, China's demand in the commercial and residential sector has rarely been examined, though it has been projected to nearly triple between 2014 and 2040 by the International Energy Agency (IEA, 2014). Moreover, this increase represents nearly half (46%) of total projected electricity growth. The U.S. Energy Information Administration (EIA) makes similar projections; however, the EIA also separates the residential and commercial sector in its International Energy Outlook (EIA, 2015). There, residential consumption alone accounts for one-third of China's electricity demand growth over the next 25 years. This implies increases in both per capita residential electricity consumption and the residential share of electricity in China in the coming decades.

This pattern is consistent with the experience of most now-developed countries. Fig. 1 relates per capita residential electricity consumption to GDP per capita in 2010 for several representative countries, along with the projected values for China through 2040 from EIA. EIA projects that per capita residential electricity consumption will more than double between 2010 and 2025. Fig. 2 gives the residential share of total electricity use for the same set of countries and shows China's projected share growing from 14% to more than 20%.

With such a large portion of projected Chinese electricity demand growth being driven by households, it is important to ask what drives these forecasts. As an example of typical forecasting approaches, the residential module in World Energy Projection System Plus (US EIA), uses only GDP, energy prices, and autonomous time trends to project energy demand. Moreover, it largely uses elasticities and trends taken from EIA's domestic U.S. model. Repeated cross-sectional data, in contrast, allows us to estimate parameters directly. We can tease out the effects of income and autonomous time trends, which are otherwise highly collinear in aggregate data. We can look for non-linear income effects, as we expect elasticities to fall at higher incomes. We can also further decompose income effects into extensive margins, as households acquire larger houses and more electricity using appliances, and intensive margins related to other, non-specific income-driven demand growth including increased use of existing homes and appliances. This kind of detail is important to improve forecasts, to understand key uncertainties, and to design policies to influence future electricity demand. The latter is particularly salient as

[1] BP Statistical Review of World Energy June 2015. http://www.bp.com/content/dam/bp/pdf/energy-economics/statistical-review-2015/bp-statistical-review-of-world-energy-2015-full-report.pdf

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

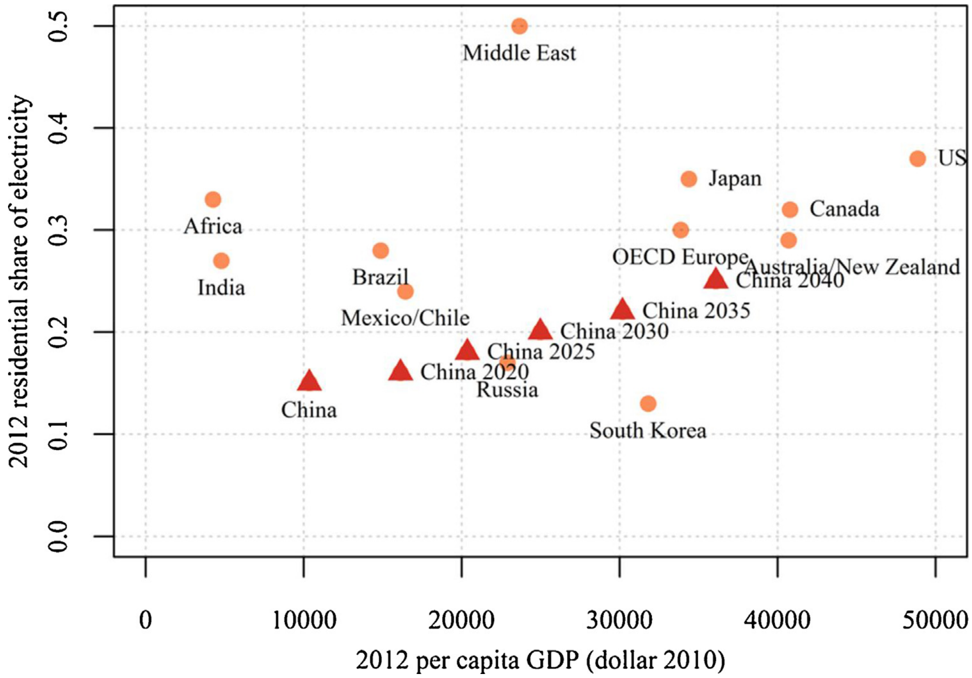

Fig. 2. Residential share of electricity versus per capita GDP in 2012, alongside projections for China.
Source: EIA, International Energy Outlook 2016

China grapples with both conventional air pollution and climate change, which would both improve with reduced household energy demand.

There are existing empirical papers on household energy demand in many countries,[2] but very few studies have been conducted in China due to the lack of household-level data, at least until recently. In the United States, O'neill and Chen (2002) focus on the role of demographic changes in aggregate US household energy demand using the Residential Energy Consumption Survey (RECS) data from 1993 to 1994. Leahy and Lyons (2010) analyze residential energy demand conditional on appliance ownership using the 2004/2005 Irish Household Budget Survey. Jones and Lomas (2015) investigate the socioeconomic and dwelling determinants of high electrical energy demand in UK households. Although some of these studies provide estimates of price and income elasticities, specifications differ and many suffer from a limited sample. This makes it difficult to credibly transfer the estimates to China given the different market and institutions. In addition, only a few studies produce forecasts using the estimated models, which is an important way to highlight the practical consequences of alternative assumptions. Our approach is to assess estimation results alongside forecast implications simultaneously.

To estimate our model and construct forecasts for China, we use a repeated cross-sectional dataset,[3] the Chinese Urban Household Survey (CUHS) collected by National Bureau of Statistics of China. This is the first paper to use such data to forecast electricity use. Two other recent studies have used these data to estimate income and price elasticities. Zhou and Teng (2013) estimated income and price elasticities of electricity demand in Sichuan province. However, they did not model demand for household appliances or consider the implications for forecasts, and did not consider the other provinces. Cao et al. (2016) estimated an energy demand system using data from all nine provinces and focused on demand choices among fuels. They similarly did not model the demand for household appliances or generate forecasts based on the results. While our parameter estimates are comparable with these studies, here our focus is building a more complete model of both appliance ownership and electricity demand. We then use that model (a revised version of the two-stage approach by Dubin and McFadden, 1984) to forecast electricity demand and to understand the drivers of future growth.

In particular, we explore the sensitivity of forecasts to a variety of alternative assumptions about prices, autonomous time trends, demographics, and income. Our central estimate projects an 85% increase in electricity demand from 2009 to 2025. Consistent with previous work in China and other countries we find a price elasticity of demand of about −0.7. To put this into perspective, if electricity prices in China were to rise by an additional 15%, consistent with the effect of a $10 per ton carbon dioxide price, household demand would decline by about 10%.

[2] See Fell et al. (2014), Espey and Espey (2004), Larsen and Nesbakken (2004), Reiss (2005), Swan and Ugursal (2009) and Alberini and Filippini (2011) for more detailed review of previous studies on residential electricity demand.
[3] Note: The CUHS has a rotational sampling each year, that is, about one-third of samples are replaced with new households.

Assumptions about autonomous time trends are considerably more important. The difference between our high and low (absent) autonomous trend models is 143% versus 43% demand growth. Meanwhile, demographic trends (other than income) have little effect. That is, income and demographic changes aside, there still appears to be an unexplained positive trend in residential electricity consumption over time in our data. If that trend continues, it would double forecast demand growth relative to estimated income effects alone. Our central estimate is based on an assumption (suggested by the data) that this autonomous trend slows over the coming decades. Given we believe the assumption of no autonomous time trend is unrealistic, our view is that the range of 85% to 143% growth is a reasonable projection for 2009–2025.

We can further decompose income-driven demand growth into components related to dwelling size and appliance adoption, as well as income alone. We find that appliance adoption drives about one-third of the growth attributable to income in our data. That is, when we consider the contribution to electricity demand growth of appliance adoption (itself driven by income growth), alongside the direct effect of income growth, appliance adoption accounts for about one-third of the total. Among appliances, air conditioners (AC) units are the most important, because of both their high energy use and increasing adoption (even as other appliances become saturated). We might label this one-third the extensive margin, with a higher stock of major appliances, and the other two-thirds the intensive margin, with other, non-specific, income-driven increases in demand, including higher use of existing appliances. Dwelling size, on the other hand, has very little independent effect.

Importantly, there are clues about potential policy targets for the intensive margin. While the number of refrigerators remains relatively unchanged in our forecast, they have the highest estimated impact on energy use. Refrigerators are estimated to increase electricity use by almost 20%. Given the doubling of refrigerator efficiency in the United States over the last 20 years, this suggests such an effort in China could reduce electricity demand by nearly 10%. This points to the potential value of energy efficiency policies, which would need to be weighed relative to the costs of such policies.

Another issue of interest is whether the distribution of income growth across different segments of the income distribution has a measurable impact on expected energy growth (Auffhammer and Wolfram 2014). We do not find the income effect to be particularly sensitive to how economic growth occurs across the different income groups. We do find that the income elasticity is falling as income rises, and project that the elasticity would reach zero when income is roughly $8200 or 57,700 RMB per capita (2010 RMB). However, ascribing more growth to richer households based on trends in the data does not substantially alter the projection.

Summarizing, our econometric analysis leaves us with a large unexplained autonomous component. While our central estimate is 85% per capita consumption growth over 2009–2025, alternative assumptions about how and whether this trend continues lower the estimate to 43% or raise it to 143%. Ruling out the possibility of no autonomous trend, our preferred range is from 85% to 143%. Despite this gap, our improved understanding of the role of appliance adoption, income, and electricity prices is helpful in thinking about future policies. Energy efficiency polices for refrigerators and modest carbon pricing, for example, might each reduce household electricity use by 10%. We owe all of these observations to our use of household level data, which allows us to separate out the effects of price, income, appliance adoption, and autonomous trends.

The remainder of the paper is organized as follows. Section 2 introduces the data with summary statistics. Section 3 explains the model framework and econometric specifications. Section 4 reports the estimation results and initial forecast results. Section 5 discusses the main forecasting results, as well as policy implications from our examination of the composition of demand growth and distribution of income growth. Section 6 concludes.

2. Data

Our objective is to explore how estimation employing microeconomic data can provide more informative projections of electricity demand for the residential sector. Forecasts based on aggregate data depend largely on extrapolating growth trends. Microeconomic data allows us to explore the relative effect of at least three distinct features observed in the cross-section: income, appliance adoption, and dwelling size. In addition, our repeated cross section over time allows us to carefully examine any remaining autonomous time trend. While these features themselves must be extrapolated, alone or jointly, their relationship with electricity demand can be estimated based on cross-sectional data.

In addition to providing greater confidence in electricity forecasts, detailed microeconomic data can provide several other benefits. First, by teasing out the relative contribution of different factors driving consumption growth, it allows us to identify the effect of key assumptions about these factors such as the persistence (or not) of autonomous time trends. Second, it allows us to simulate the impact of alternative assumptions about the distribution of aggregate economic growth. That is, how does electricity consumption vary if growth is broad-based versus being more concentrated among the wealthy, as it has tended to be in the past. Finally, a more detailed model of residential energy use, including prices, appliances and dwelling size, allows us to assess the likely impacts of various policies, including electricity pricing and energy efficiency policies.

In this section, we describe the data sets we use and the summary statistics of variables of interest, followed by a discussion of the model and specification in the next section.

2.1. Chinese urban household survey (CUHS)

This paper uses the cross-sectional Chinese Urban Household Survey (CUHS) data collected by the National Bureau of Statistics (NBS) from 2002 to 2009. The survey design was changed in 2007, but the change has almost no impact on the variables chosen for our analysis.[4] Households with *hukou* (registered residency status in a particular area) are included[5]; households without *hukou* would be counted if they live in the city for more than six months. The total number of households from 31 provinces surveyed is around 60,000 each year, of which about 12,000 to 16,000 households per year are from 153 cities in our 9 provinces. The 9 regions are Anhui, Beijing, Guangdong, Hubei, Liaoning, Shaanxi, Sichuan, Zhejiang, Gansu, covering all parts of the country except the sparse far west.[6] The survey employs stratified multistage random sampling to guarantee a representative sample at each level. The 9 provinces that we have access to are not randomly selected; they reflect a restricted sample provided by NBS. Therefore, we compare the 9-province averages to the national averages reported by the NBS to show the similarity. One measure is the number of appliances owned per 100 urban residents. The number of AC's per 100 urban residents is 68 in our sample compared to 51 for the national average in 2002. However, the difference narrowed to less than 10 by 2009. The numbers for washers, refrigerators, computers, and microwave are also slightly higher in the 9-province sample. These nine provinces are slightly more developed compared to the rest at the beginning of the sample period, but the gap diminishes over time.

The CUHS data is very comprehensive in coverage, including socioeconomic features, and detailed information on consumption expenditures. In this study, we mainly focus on electricity consumption, relevant socioeconomic variables, housing features and the stock of appliances. Data on household features (e.g., household size, dwelling size, electricity consumption) and data on household head characteristics (e.g., age and education level) are contained in two separate datasets for each of the eight years. We merged the two datasets by household identity.

Electricity prices are calculated from electricity expenditures (in RMB) and quantities (in kWh).[7] There are ample variations in the electricity prices across provinces and across cities within provinces. The calculated prices have relatively small variations across different districts in Beijing (the same pricing policies apply to all districts) but we observe larger variations across cities in other regions (as pricing rules usually differ from city to city). The calculated electricity prices in Guangdong are the highest, ranging from 0.6 to 1 RMB/kWh, while the prices in other eight regions are lower, with the majority falling in the 0.4–0.6 RMB/kWh range. Given potential endogeneity concerns about using consumption to compute the price, we also estimate our model using the average electricity price at the city level as part of our sensitivity analysis.

The sample provided by NBS includes cross-sectional sampling weights. We use these weights in our estimation to ensure that our parameters are appropriate for projecting population averages. Moreover, the weights for each year are normalized so that the total weight given to each year in the sample is the same.[8] That is, the NBS sample size varied from about 12,500 in 2002 to 16,500 in 2005–2009. The years with a smaller sample are weighted upwards to ensure that years with more observations do not have more influence on the parameters estimates. We cleaned the data through several steps (see Appendix A).

Real RMB values (for income and the electricity price) are generated from the original, nominal survey data using both the provincial-level price index for a common basket calculated by Loren Brandt and Carsten Holz[9] and the official CPI data. To capture the provincial differences in price levels that is not in the official CPI, Brandt and Holz calculated the urban basket cost for different regions. We use the urban basket cost from their calculations for the base year 2002 as our price index for the first year (the price index of Beijing is normalized to 100 in 2002). The price indexes for 2003–2009 are then generated using the official provincial CPI inflation rate.

Finally, to examine the impact of weather on electricity consumption and better model the effect of AC units on demand, we combined the survey data with data on Heating Degree Days (HDD) and Cooling Degree Days (CDD). City-level HDD or CDD data are calculated based on the weather data from the *Global Summary of the Day* from the *National Oceanic and Atmospheric Administration* (NOAA), which has 50 weather observation points close to or in the 153 cities in our sample.[10] Temperatures are computed for each city based on the nearest observation point. To calculate HDD and CDD, we use 65 °F (18 °C) and 78.8 °F (26 °C) as the reference values for heating and cooling respectively. Although many previous studies in Europe and the United States use 71.6 °F (22 °C) as the reference value for cooling, we follow the recent studies (e.g., Shi

[4] Two types of refrigerators for 2002–2006 are combined into one variable for 2007–2009.

[5] Note: migrant workers from rural areas are not included.

[6] The CUHS data for these nine provinces were provided by the China Data Center, Tsinghua University.

[7] The calculated electricity price is the price averaged over the year for a household. Increasing empirical evidence has supported that consumers are more likely to respond to average prices (Fell et al., 2014). Also, with a double-log form, elasticity estimates for marginal and average rates are quite comparable (Halvorsen, 1975). Note that the tier-pricing policies have not yet been implemented during 2002–2009 hence households would not be able to influence the average prices by changing consumption.

[8] More specifically, if qs_{it} represents the NBS sampling weight for a given household i in year t, then the normalized weight $q\tilde{s}_{it} = qs_{it}/\sum qs_{it} \times 10,000$. After normalization, the sum of the weights in each year is 10,000. Here, the arbitrary choice of "10,000" to normalize the population each year does not affect estimation.

[9] See Brandt and Holz (2006). More recent data is made available at http://ihome.ust.hk/~socholz/SpatialDeflators.html

[10] The data is described at https://data.noaa.gov/dataset/global-surface-summary-of-the-day-gsod.

Table 1
Summary statistics of household data in CUHS, 2002–2009.

N = 122,252	variable	mean	weighted mean	sd	min	max	%change*
Electricity consumption (kWh/year)	*q*	1430	1484	1047	2	24,740	7.4%
Real electricity price (RMB/kWh)	*p*	0.448	0.460	0.121	0.041	1.5	−0.9%
Real income (RMB/year)*	*y*	22,968	22,957	20,758	203	947,186	6.0%
Demeaned real income	*y_c*	1	1	0.900	0.010	41.080	6.0%
Demographic features							
Dwelling size (square meters)	*dwelling_sz*	78.545	82.215	36.085	5	300	2.2%
Rent (RMB/square meter/month)	*rent*	7.094	6.783	9.417	0	224	106%
Ownership (=1 if rent)	*ownership*	0.127	0.124	0.333	0	1	−0.9%
Household size	*household_sz*	2.915	2.961	0.817	1	13	−0.9%
Age of household head >50 (0/1)	*age_over50*	0.410	0.396	0.492	0	1	1.0%
Education of household head (0/1)							
Primary or below	*edu_prim*	0.072	0.069	0.259	0	1	0.1%
Middle school or equivalent	*edu_mid*	0.279	0.265	0.448	0	1	−0.2%
High school or equivalent	*edu_high*	0.350	0.364	0.477	0	1	−0.7%
College level or above	*edu_coll*	0.299	0.302	0.458	0	1	0.7%
Heating type (0/1)							
Heating by air-conditioner	*heat_ac*	0.274	0.262	0.446	0	1	3.0%
Heating by gas or water	*heat_gaswater*	0.331	0.234	0.471	0	1	0.2%
Heating by other	*heat_other*	0.080	0.059	0.271	0	1	−0.5%
None	*heat_none*	0.315	0.445	0.464	0	1	−0.6%
Appliances (0/1/≥2)							
Refrigerator	*fridge*	0.962	0.962	0.364	0	2	0.0%
Computer	*comp*	0.469	0.465	0.558	0	2	14.3%
Microwave oven	*mwave*	0.488	0.488	0.507	0	2	7.6%
Waterheater for shower	*wheater*	0.784	0.812	0.479	0	2	2.7%
Washing machine	*washer*	0.953	0.960	0.316	0	2	0.2%
Air conditioner (0/1/2/3/>=4)	*ac*	0.894	0.961	0.970	0	4	8.0%
Weather							
Heating degree days	*hdd*	3739	3221	1892	273	7833	2.4%
Cooling degree days	*cdd*	285	336	224	0	919	0.7%

Note: For household real income, we use consumption expenditure instead of the reported income. There is no life-time income estimation using cohort studies as in US so we simply use consumption expenditure as a proxy for life-time income in our study. For continuous variables and the appliances, the '%change' is calculated as the average annual growth rate for 2002–2009. For indicators (0/1), the '%change' is calculated as the mean annual change for 2002–2009.

et al., 2016) on China to reflect local habits, building characteristics and climate conditions in China. More specifically, HDD and CDD for city i and year t are calculated as follows:

$$\begin{cases} HDD_{i,t} = \sum_{d=1}^{365} \left\{ \max\left(65 - T_{i,t,d}, 0\right) \right\} \\ CDD_{i,t} = \sum_{d=1}^{365} \left\{ \max\left(T_{i,t,d} - 78.8, 0\right) \right\} \end{cases} \tag{1}$$

where $T_{i,t,d}$ is the mean temperature on day d of year t in city i.

2.2. Descriptive statistics

Table 1 shows the descriptive statistics for the variables used in the econometric estimation. The sample size is 122,252 households (spread across 8 years). The data used in this study and Cao et al. (2016) are, to our knowledge, the largest and most representative micro data set ever used for household energy research in China. The first four rows include key variables—electricity use, price, and income. This is followed by demographic features, education of household head, heating type, appliances ownership and weather. In addition to the standard summary statistics (mean, standard deviation, minimum and maximum), the table includes the weighted mean estimated using the population weights in the data (column 4). We also calculate the annual percentage change of the within-year average for each variable over the 8-year period (the last column).

In weighted average terms, electricity consumption per household is 1484 kWh, real income, proxied by total consumption expenditures is 23,000 RMB (roughly $3500),[11] and dwelling size is 82 square meters. The average per capita income would be roughly one-third, given the average household size of about 3. Among appliances, the ownership rates of refrig-

[11] In comparison, the average disposable income is 29,789 RMB.

erators and washing machines are over 95%, followed by an AC penetration rate of 89%, and computer of almost 50%. The average heating degree days and cooling degree days are 3200 and 340 respectively.

Looking at trends over time, real income and electricity consumption increased at an average annual rate of 6.0% and 7.4%, respectively, which is consistent with the increasing per capita GDP and electricity demand at the macro level. Real electricity prices declined slightly. Among demographic features, household size decreased while dwelling size increased by 2.2% annually. Over the sample period, education levels generally rose with an increase in the number of household heads with a college education or above. Among appliances, the ownership of computers, AC's and microwave ovens grew rapidly while refrigerators and washing machines are approaching saturation.

3. Method

Electricity consumption is associated with the use of a range of household electrical devices that are operated for many years, if not decades. Hence, understanding and modeling the ownership of such appliances is critical in analyzing electricity demand in the residential sector, particularly if one's interest is ultimately related to understanding the link to income growth, to electricity prices, and to policy interventions. In addition to appliances, larger dwelling size would increase electricity consumption through higher heating or cooling demand.

Fisher and Kaysen (1962) proposed a two-stage framework that informs our modeling choice. They model consumption in the short-run as a function of income and price, given the appliance stocks. In a separate saturation model, they use population, expected income, marriages, expected energy prices and the number of wired households to explain the appliance stocks. The empirical results from their paper are constrained by the quality of the appliances data, but the two-stage framework has been found to be an improvement in results relative to a static model (Taylor et al., 1984). More recently, Leahy and Lyons (2010) use the two-stage framework to estimate a logit model investigating the determinants of household appliances and OLS models to explain household energy use conditional on appliance ownership.

Expanding on the above framework, we assume that households face a three-stage sequential and independent decision process to make each step tractable. In stage I, they choose *the size of house* to rent or own. In stage II, conditional on the dwelling size chosen, they choose *the number of appliances for each type*. In stage III, conditional on the dwelling size and the appliances chosen, households decide *how much electricity to consume*. Each stage is a function of income and other demographic variables, possible time trends, and previous stage(s) dependent variable(s). The last two stages include the electricity price. We adopt this three-stage framework mainly to take into account the non-linearity in the adoption of appliances and to decompose the sources of growth of electricity demand. In particular, the three-stage model allows us to compare these driving forces: dwelling size, appliance ownership, and utilization rate given dwelling size and appliances. This then allows us to both more confidently forecast demand and consider the magnitude of possible policy interventions.

Income is a main factor of interest so we tested different functional forms using linear, quadratic and cubic functions. We choose the quadratic form of income based on both economic intuition and statistical significance (particularly for the last stage). A negative quadratic term is consistent with diminishing marginal income effects. As income rises, a one percent increase in income would lead to a smaller increase in dwelling size and electricity consumption. For appliances, S-shape curves have been documented (McNeil and Letschert, 2010), where households start adopting appliances after passing an income threshold, first experiencing a speed-up in adoption and then a deceleration. Our ordered logit model imposes such behavior without a quadratic income term; however, we include it in order to provide a more flexible income-adoption relationship and for consistency with the other two stages.

As we do not have panel data at the household level, we apply the three stages to the pooled repeated cross-section of households. For each stage, we estimate a relatively simple reduced-form model. We use ordinary least squares (OLS) for dwelling size, an ordered logit (Ologit) model for appliances, and OLS for electricity consumption. This allows flexible, transparent estimation of the parameters of interest as well as projection using available income forecasts. We describe additional details below.

3.1. Dwelling size model

Housing demand is obviously not a simple choice of structure and size, but is also a location decision, where the location attributes of interest include proximity to work, schools and amenities, and environmental quality among many others. The price of a particular dwelling depends on all these characteristics. The survey data provides a per-square-meter housing price, including a rental equivalent for owner-occupied units, which we use.

Given these data constraints, in stage I, dwelling size is modeled as a continuous choice based on income, price represented by rent, household size, and other demographic features. For each household from a given year:

$$\begin{aligned}\ln\ (dwelling\_sz_{it}) &= \alpha + \beta_1 \ \ln\ (rent_{it}) + \beta_2\ \ln\ (y\_c_{it}) + \beta_3\ [\ln\ (y\_c_{it})]^2 + \beta_4 household_{sz\,it} + \\ &\beta_5 age_{over50it} + \beta_6 edu\_mid_{it} + \beta_7 edu\_high_{it} + \beta_8 edu\_coll_{it} + \beta_9 ownership_{it} + \delta_{region} + g_k(t, b_1, b_2, \delta_{year}) + \epsilon \qquad (2)\end{aligned}$$

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

where the β's are parameters to be estimated; *dwelling_sz* is the dwelling size; $rent_{it}$ is the monthly rent for the house or the equivalent rent if the house is owned[12]; *y_c* is the demeaned real income represented by the real consumption expenditure[13]; *household_sz* is the household size (number of people); *age_over*50 is an indicator variable for households with a head older than 50 years old; *ownership* is an indicator that equals 1 for rental houses; and δ_{region} are fixed effects for each of the 9 provinces. The education level of household head is captured by three indicators: middle school (*edu_mid*), high school(*edu_high*), college and above(*edu_coll*), all relative to primary education (which is omitted).

The general form $g_k\left(t, b_1, b_2, \delta_{year}\right)$ represents the time trend: t is defined by $t = year - 2002$; b_1 and b_2 are parameters; and δ_{year} are possible year indicators for 2002–2009. We consider three different specifications of $g_k(.)$ (k = A, B, C) to capture different possibilities for how time trends may continue into the future. In general, there is an upward trend from year to year in the data that is not explained by other demographic variables. The alternative specifications, all of which allow for this historic pattern, have little effect on the estimation of other parameters but affect future forecasts significantly. Applying different time trends allows us to examine the range of possible trajectories. Specifically, our first specification simply uses year fixed effects, which we then project forward based on the fixed, final year value, the second one uses a non-linear, slowing time trend, and the third specification uses a linear time trend:

$$\begin{aligned} g_A(t, b_1, b_2, \delta_{year}) &= \delta_{year} \\ g_B(t, b_1, b_2, \delta_{year}) &= -b_1 e^{-b_2 t} \\ g_C(t, b_1, b_2, \delta_{year}) &= b_1 t \end{aligned} \tag{3}$$

A variety of individual effects is embedded in both demographic variables and the error term. In the period before economic reform, housing was allocated by the place of employment, which complicates the dwelling choice. In the 1990s, the state-owned or enterprise-owned housing was privatized. Some organizations, especially government agencies, may still build and sell subsidized units to their staff. Employment, in turn, is likely determined in part by education and age. We should also note that urban China today is unusual, compared to developed countries, in that there is a very high ownership rate; less than 20% rent their housing. The consumption expenditures we employ to represent household income, therefore, do not include the market value of housing services for a large portion of the data; this is a potential limitation to our measure of income.

3.2. Appliance stock model

In stage II, we model each of six appliances separately. For air-conditioners (AC), the choice is categorized into owning 0, 1, 2, 3, or 4+ units (represented by 4 hereafter). For the rest, including washing machine, refrigerator, microwave, computer and water-heater, the choices are owning 0, 1 or 2+ (represented by 2 hereafter). Based on Dubin and McFadden (1984), we model the appliance choices in a logit framework. As the possibility of the fact that multiple appliances are interrelated, the ordered logit is most suitable to capture the ordinal scale of the dependent variables. That is, someone more likely to purchase two rather than fewer devices is also more likely to purchase one or more devices rather than zero.

We illustrate the model set up for AC for each household i for a given year t; the other appliances with outcomes of up to 2 units follow similarly. We observe the outcome of AC ownership in an ordinal scale of 1, 2, 3, 4. We assume that there is an underlying variable measuring the utility from AC, denoted $\tilde{AC}$. Whether household i owns a particular count of AC units depends on whether the underlying driver $\tilde{AC}_i$ passes particular thresholds. More formally, for a household in a given year:

$$\begin{aligned} \mathrm{AC_{it}} &= 0 \; if \; \tilde{AC}_{it} \le \kappa_1, \\ \mathrm{AC_{it}} &= j \; if \; \kappa_j < \tilde{AC}_{it} \le \kappa_{j+1}, for \; j = 1, 2, 3 \\ \mathrm{AC_{it}} &= 4 \; if \; \tilde{AC}_{it} > \kappa_4, \end{aligned} \tag{4}$$

where κ represents "cuts" or thresholds to be estimated. The continuous latent variable $\tilde{AC}$ is modeled as a linear regression on covariates:

$$\tilde{AC}_{\mathrm{it}} = X_{it}\beta + \varepsilon_i \tag{5}$$

where β is a vector of coefficients; X_{it} is a vector of covariates, including all the variables contained in the dwelling size regression (Eq. (2)), demeaned dwelling size and demeaned dwelling size squared, and electricity price. We also include cooling degree days (*cdd*) and heating degree days (*hdd*) in the model for AC, but not for the other appliances. One limitation of our model is that we do not have information on the price of appliances. Through the province and year fixed effects,

[12] Note: The average rent is surprisingly low in the sample, but the trend over time reflects the housing boom in 2005–2007. Assuming the variable is recorded consistently across households, we believe it is better to include it to avoid omitted variable bias.

[13] Note: we use consumption expenditure as a proxy for life-time income in our study.

we implicitly control provincial-specific and year-specific price differences, but not the price differences associated with different size, brand, quality, and energy efficiency.

The random disturbance term ε_i is assumed to follow a standard logistic distribution. Thus, the probability for AC taking on a particular value is given as follows:

$$\begin{aligned} &\mathrm{P}(\mathrm{AC_{it}}=0)=\frac{1}{1+\exp(X_{it}\beta-\kappa_1)}\ ;\\ &\mathrm{P}(\mathrm{AC_{it}}=\mathrm{j})=\frac{1}{1+\exp(X_{it}\beta-\kappa_{j+1})}-\frac{1}{1+\exp(X_{it}\beta-\kappa_j)}, for\ \ j=1,2,3\\ &\mathrm{P}(\mathrm{AC_{it}}=4)=1-\frac{1}{1+\exp(X_{it}\beta-\kappa_4)} \end{aligned} \tag{6}$$

For each appliance, the parameters including both the β's and κ's can then be estimated by maximizing the log-likelihood function:

$$\log\ \ L(\beta)=\sum\{I[appliance_{it}=j]\cdot \mathrm{P}(\mathrm{appliance_{it}}=j)\},\ \begin{cases} j\in\{0,1,2,3,4\}\ for\ ACs\\ j\in\{0,1,2\} for\ \ other\ durables\end{cases} \tag{7}$$

where $I\,[appliance_{it}=j]$ is the indicator function that equals 1 when the number of appliances equates j for a household i in year t.

3.3. Electricity consumption model

In stage III, based on the previous studies and examination of different models, the regression model for electricity consumption conditional on appliances by household i in year t is specified as:

$$\begin{aligned} &\ln\ \ (ec_{it})=\alpha+\beta_1\ \ \ln\ \ (y\_c_{it})+\beta_2\ \ [\ln\ \ (y\_c_{it})]^2+\beta_3\ \ \ln\ \ (p_{it})+\beta_4\ \ \ln\ \ (dwelling\_sz_{it})+\beta_5 household\_sz_{it}+\\ &\beta_6 age\_over50_{it}+\mathrm{B}_7 hhedu_{it}+\mathrm{B}_8 appliances_{it}+\beta_9 hdd_{ct}+\beta_{10} cdd_{ct}+\mathrm{B}_{11} heating_{it}+\mathrm{B}_{12} hdd_{ct}\cdot heating_{it}+\\ &\beta_{13} cdd_{ct}\cdot AC_{it}+g_k(t,b_1,b_2,\delta_{year})+\delta_{region}+\epsilon \end{aligned} \tag{8}$$

Where ec is household electricity consumption (in kWh), the β's are parameters and **B's** are vectors of parameters. The income, price,[14] dwelling size and demographic variables are the same as in the first two stages (with education indicators collapsed into the vector *hhedu*). In addition, *appliances* represents the number of specific appliances owned, including washing machine, refrigerator, computer, microwave oven, water heater for shower and AC; *hdd* and *cdd* are heating degree days and cooling degree days respectively; *heating* has the indicators for heating types, including heating through an air conditioning unit (*heat_ac*) heating by gas or water (*heat_gaswater*), and heating by other (*heat_other*). The category "no heating" is omitted as the reference group. We also include interactions between heating degree days and heating types ($hdd{\cdot}heating$), and between cooling degree days and ACs($cdd\cdot AC$). Note that *hdd* and *cdd* vary by city (using subscript c) and by year. As in the dwelling size model, the specification in Eq. (8) includes the term, $g_k\left(t,b_1,b_2,\delta_{year}\right)$, for the three options for trends – year fixed effects, non-linear time trend and linear time trend.

4. Estimation and forecast results

We now turn to estimating the model and generating forecasts based on it. We discuss each component of the three-stage approach in turn in this section, and then consider the broader observations in the next section. For each forecast, we include both the time trends and income growth. Future per capita income growth is calculated from the projections of GDP and population in the IEA's World Energy Model 2015.[15] Real GDP is assumed to grow at 9.9% per year for 2010–2013, 6.4% for 2014–2020 and 5.3% for 2021–2025 while population is assumed to grow at 0.4% until 2025 (WEM, 2015).[16] Hence, the per capita GDP growth rate is 9.5% for 2010–2012, 6.0% for 2013–2020 and 4.9% for 2021–2025. This amounts to a 165% increase from 2009 to 2025, or a roughly one-unit increase in logged per capita income. We use the forecasted rate by IEA mainly to maintain consistency of income assumption over the whole period of 2010–2025.

To construct forecasts, we use the model to predict new values for each of the modeled variables (dwelling size, appliance ownership, and electricity consumption) for each of the households observed in 2009 based on per capita GDP growth[17] and the specified time trend. We do this in sequence so that forecast dwelling size influences appliance ownership, and dwelling

[14] We use the electricity price calculated at household-level for the main specification. This may cause endogeneity issue if the calculated price is correlated with other uncontrolled variables. As an alternative, we use the average price at the city level which would not be correlated with household-level unobservables. The results do not change significantly.
[15] http://www.worldenergyoutlook.org/media/weowebsite/2015/WEM_Documentation_WEO2015.pdf.
[16] Note: As both the urban and rural areas are included by the WEM forecasts, we assume that the urban households follow the same GDP growth rate as the rural households.
[17] Note that the forecasts assume no change in electricity prices or degree of urbanization.

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

Table 2
OLS estimation results for dwelling size (dependent variable = ln (*dwewlling_sz*)).

	Model A	Model B	Model C
lnrent	−0.0598***	−0.0371***	−0.0325***
	(0.0133)	(0.00918)	(0.00858)
lny_c	0.0950***	0.0899***	0.0887***
	(0.0153)	(0.0161)	(0.0163)
lny_c^2	−0.000375	−0.00165	−0.00197
	(0.00573)	(0.00598)	(0.00610)
household_sz	0.0591***	0.0602***	0.0603***
	(0.00750)	(0.00765)	(0.00771)
age_over50	0.0480***	0.0494***	0.0501***
	(0.00957)	(0.00965)	(0.00970)
edu_mid	0.0265*	0.0263*	0.0269*
	(0.0110)	(0.0109)	(0.0109)
edu_high	0.0734***	0.0736***	0.0748***
	(0.0131)	(0.0132)	(0.0132)
edu_coll	0.178***	0.178***	0.179***
	(0.0155)	(0.0157)	(0.0157)
ownership	−0.481***	−0.464***	−0.461***
	(0.0256)	(0.0238)	(0.0235)
α	4.159***	4.157***	3.964***
	(0.0594)	(0.0775)	(0.0590)
Time trend coefficients			
b_1		0.207***	0.0223***
		(0.0485)	(0.00374)
b_2		0.235**	
		(0.0835)	
year indicators	*YES*	*NO*	*NO*
regional indicators	*YES*	*YES*	*YES*
N	110319	110319	110319
adj. R^2	0.281	0.277	0.276

Note: Clustered standard errors in parentheses, _c stands for demeaned variable.
F-test suggests that b_1 and b_2 are jointly significant at 1% significant level.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

size and appliance ownership in turn affect electricity consumption. The households are then averaged using the 2009 cross-sectional NBS dataset weights to generate an estimated population forecast of average urban household electricity demand.

4.1. Dwelling size results

As discussed above, our model for dwelling size (Eq. (2)) includes income, other socioeconomic variables and alternative time trends. Table 2 reports the estimation results for the dwelling size model using ordinary least squares (OLS). Models A-C present alternative results for different assumptions about the autonomous time trend. Model A includes indicator variables for each year; Model B includes an exponential time trend of the form $-b_1e^{-b_2t}$ and Model C includes a linear time trend. The coefficients for all other explanatory variables have minor differences among the three specifications, suggesting that using a non-linear trend, linear trend, or year indicators does not significantly affect the parameter estimates on the other regressors.

The impacts of the non-income variables are as follows. Higher rent reduces the dwelling size. Compared to houses that are owned, the dwelling size is 46% smaller for rental houses. An older (over 50) head of household is associated with a 6% higher dwelling size. Higher education is also associated with greater dwelling size, with as much as an 18% larger dwelling for college-educated versus primary school (the reference group). Finally, each additional household member is associated with an estimated 6% larger dwelling size. These variables are all held fixed based on their distribution in 2009 in the forecasts. We later show that our results are not particularly sensitive to alternative trends in demographic variables.

Income is the variable of greatest interest, as we use income projections (in part) to drive dwelling size forecasts. While the linear term suggests a statistically significant elasticity of 0.09—that is, a 100% increase in income raises dwelling size by about 9%—the quadratic term is not statistically significant. Moreover, the point estimate of the quadratic term suggests that dwelling size does not "peak" until income is 400 times higher than the 2009 income level; thus there is little evidence that growth in dwelling size is slowing yet as a function of rising incomes. We keep this term in the model, however, as

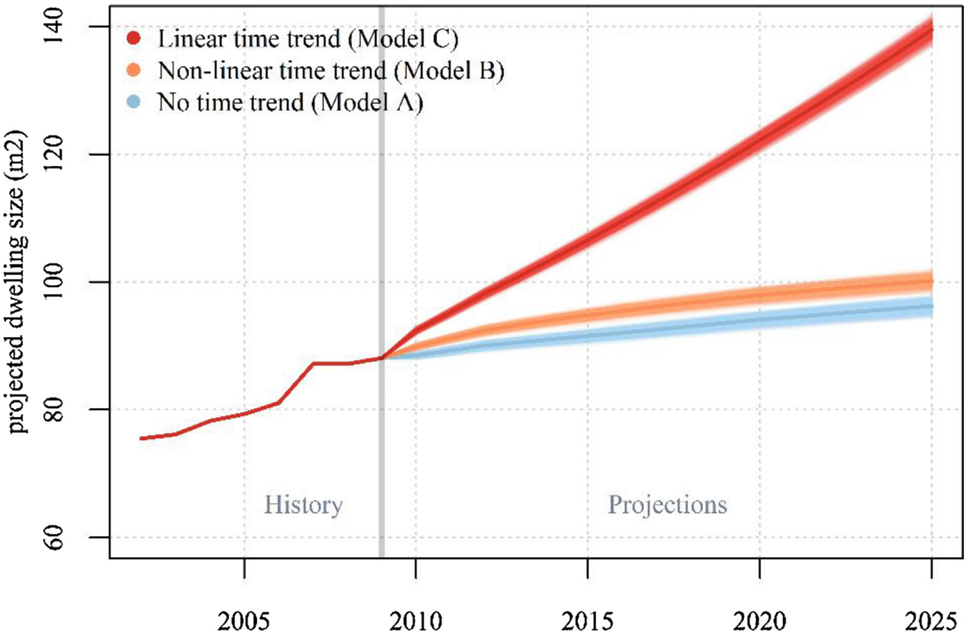

Fig. 3. Mean dwelling size (in square meters) projected from 2009 using alternative time trends. Shaded regions indicate variation based on uncertain demographics.

it has the expected sign and we believe there should be some slowing in the otherwise linear relationship. Based on this estimated relationship, over the forecast horizon dwelling size will rise 10% based on forecast income growth.[18]

The other item of interest is the choice among alternative time trends. The model fit is virtually identical with year indicators in Model A versus Models B and C with parameterized time trends over the eight years in the sample (adjusted R^2 of 0.28). However, for forecasting purposes, we prefer a parameterized time trend; otherwise we have to assume the year fixed effect in the last year persists forever. With the linear time trend model, dwelling size would rise 58% over the forecast horizon. Finally, the parameters b_1 and b_2 of the non-linear time trend in Model B are both significant. The point estimates suggest that, from 2009 levels, dwelling size would increase by 14% over the 2025 forecast horizon and eventually achieve a 21% increase asymptotically (holding other variables, including income, constant). Note that we have not forecast housing prices explicitly in this projection. With an estimated price elasticity of roughly −0.05 across models A-C, this could amount to 5 percent decrease in dwelling size if housing prices doubled.

Perhaps the most useful way to compare the time trends is to see how they influence the forecast. Fig. 3 shows the three forecasts of dwelling size based on the three different time trend models and the per capita income growth projection described above. The shading in the figure represents sampling error in the demographic variables used to forecast the model.[19] This sampling error is small compared to the notable effect of alternative time trends on the trajectory over the forecast period. Model A predicts a 9% increase; Model B predicts a 14% increase and Model C predicts a 58% increase from 2009 to 2025. That is, there is significantly higher growth in dwelling size if we extrapolate linear time trends versus keeping them fixed at 2009 levels. This is true even after controlling for income growth and rent. As we will show later, the difference in dwelling size projections, although large in and of themselves, alter the ultimate projection of electricity demand only slightly.

4.2. Appliance stock results

Table 3 reports the estimation results for the choice of appliances. All six appliances are modeled through the ordered logit as a function of the same variables as the dwelling size model, plus dwelling size and the electricity price. In an ordered logit, the probability of being at or above a particular count *j* for the indicating appliance is $\exp\left(X\beta-\kappa_j\right)/\exp\left(1+X\beta-\kappa_j\right)$ where κ_j is the cut. The lower part of Table 3 shows the cuts, which are not exponentiated. The cuts are all significantly

[18] As mentioned in the method section, we use consumption expenditure as the proxy for lifetime income. One concern is that purchasing houses would affect consumption expenditure, causing an endogeneity issue. To address this issue, we estimate the same specification using the total household income. The income elasticity increases slightly, leading to upward-adjustment of roughly 2% in dwelling size over the forecast horizon.

[19] We resample households from the 2009 cross-section data with a replacement before applying the income growth and time trend assumptions to each household, which are then averaged to produce a forecast. We then conduct this exercise 1000 times to produce the shaded region.

Table 3
Ordered logit estimation results for appliances ownership (odds ratio).

	Refrigerator	Washing Machine	Computer	Microwave	Water-heater	AC
lny_c	3.099***	2.563***	4.588***	3.654***	2.670***	5.125***
	(0.269)	(0.128)	(0.235)	(0.162)	(0.133)	(0.297)
lny_c^2	0.854***	0.937*	0.854***	0.819***	0.916*	0.903*
	(0.0324)	(0.0247)	(0.0302)	(0.0291)	(0.0389)	(0.0459)
lnp	0.418*	0.637	0.434**	0.651	0.601	0.479
	(0.186)	(0.210)	(0.111)	(0.182)	(0.278)	(0.271)
lndwelling_sz_c	1.458**	1.735***	1.399***	1.298**	2.722***	2.304***
	(0.184)	(0.179)	(0.115)	(0.118)	(0.447)	(0.241)
lndwelling_sz_c^2	0.776**	0.838	0.850*	0.790***	0.742*	0.855
	(0.0707)	(0.0775)	(0.0634)	(0.0559)	(0.0972)	(0.0805)
household_sz	0.958	1.123***	0.967	0.813***	0.821***	0.815***
	(0.0365)	(0.0282)	(0.0252)	(0.0194)	(0.0214)	(0.0235)
age_over50	1.305***	1.196***	0.919	1.194***	1.059	1.275***
	(0.0803)	(0.0637)	(0.0445)	(0.0635)	(0.0472)	(0.0625)
ownership	0.617***	0.643***	0.759***	0.715***	0.559***	0.563***
	(0.0748)	(0.0513)	(0.0459)	(0.0492)	(0.0374)	(0.0581)
edu_mid	1.435***	1.626***	1.542***	1.523***	1.385***	1.437***
	(0.132)	(0.126)	(0.0960)	(0.0881)	(0.0758)	(0.0750)
edu_high	1.751***	2.087***	2.157***	2.133***	1.804***	1.864***
	(0.162)	(0.155)	(0.128)	(0.124)	(0.104)	(0.121)
edu_coll	1.830***	2.506***	3.693***	2.961***	2.185***	2.544***
	(0.176)	(0.182)	(0.240)	(0.179)	(0.160)	(0.190)
cdd (in 1000)						6.596***
						(3.702)
hdd (in 1000)						1.036
						(0.101)
cut1	0.293*	0.140***	4.164***	2.506	0.504	49.13***
	(0.149)	(0.0488)	(1.382)	(1.242)	(0.211)	(42.54)
cut2	138.6***	127.3***	255.3***	1502.3***	170.4***	593.0***
	(72.92)	(43.54)	(85.01)	(765.9)	(78.51)	(526.9)
cut3						3227.1***
						(2915.8)
cut4						18653.3***
						(17237.2)
year indicators	*YES*	*YES*	*YES*	*YES*	*YES*	*YES*
regional indicators	*YES*	*YES*	*YES*	*YES*	*YES*	*YES*
N	122252	122252	122252	122252	122252	122252

Exponentiated coefficients; Clustered standard errors in parentheses; cuts correspond to coefficients, not exponetiated.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

distinct from each other based on their standard errors, suggesting that assuming an underlying rank of utility level for the choices is meaningful and ordered logit models are suitable in this case.

The upper part of Table 3 shows the coefficients in the form of the odds ratio, i.e., the exponentiated coefficients. For example, the 0.418 estimate for *lnp* associated with refrigerators implies that the odds ratios P(fridge ≥ 1)/P(fridge < 1) and P(fridge ≥ 2)/P(fridge < 2) would both be multiplied by a factor of 0.418 (e.g., the ratio would decline) as *lnp* increases by 1 unit. Generally, all the coefficients on *lnp*, the logged electricity price, are around one-half—indicating that higher electricity prices lower the probability of higher numbers of appliances—though most are not statistically significant.

Except for household size and home ownership, all other socioeconomic variables (age, education, and income) have statistically significant, positive (>1) effects. A college degree raises the odds of a higher count of computers by a factor of almost four; computers are the one item where an older head of household does not increase the odds ratio. Income initially raises the likelihood of more appliances but, with the squared income term less than one, eventually lowers the likelihood. However, by simulation, we know that this peak does not occur until average income is at least 20 times the 2009 level for microwave ovens, 30 times for refrigerator and water-heater, 60 times for computer and 30,000 times for washing machine and AC.[20] Given income rises only by a factor of 2.75 in our 2010–2025 forecast, these peaking events remain far beyond the 2025 horizon.

Dwelling size also has a statistically significantly positive impact on the adoption of appliances. Here, the quadratic terms suggest a peaking for computer, refrigerator, and microwave counts when dwelling size has roughly doubled from the current mean. Given dwelling size increases by only 60% by 2025 under the most aggressive trend assumption, this also

[20] The log of this multiple can also be computed based on ln(linear term) ÷ (2 ln(squared term)) as the coefficients have been exponentiated.

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

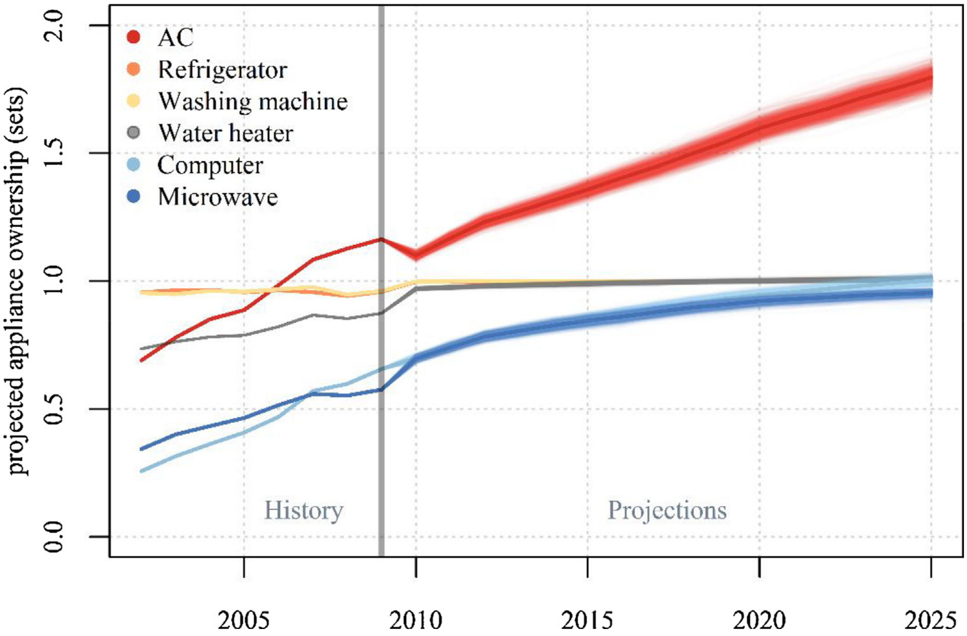

Fig. 4. Mean appliances ownership (in sets) projected from 2009. Shaded regions indicate variation based on uncertain demographics.

remains beyond our forecast horizon. Holding the dwelling size fixed, appliances ownership is significantly lower for the rental houses.

Only household size (aside from electricity price and home ownership) lowers the expected appliance count for microwaves, water heaters, and AC. This presumably reflects the effect, holding income and dwelling size constant, of a reduction of disposable income available for appliances when there are more household members. An example is the expenses for children. This effect does not affect refrigerator or computer expected counts and the expected washing machine count goes up, which may be viewed as a greater necessity in a larger household.

We include cooling degree days as a predictor of AC adoption. One thousand additional degree days (roughly 3 °C warmer every day of the year) raises the odds ratio by a factor of 6. Higher heating degree days also increases the likelihood of owning an AC, but not significantly. This is because northern China mainly uses central heating and southern China rarely needs AC for heating. Therefore, the variations in hdd within central China largely identifies the effect, based on our inclusion of region fixed effects. One limitation of the appliance models is the lack of specific equipment attributes (e.g., energy efficiency), which also play a significant role in electricity consumption.

We use these estimated adoption models to predict appliance ownership in Fig. 4. To construct this figure, we use the projected income per capita growth described above, along with the dwelling size prediction based on Model B, to predict the probabilities of different counts for each household in the sample. We take the expected value for each household in each year based on these probabilities and construct a population average using the sample weights.

It is evident that even though we allow for ownership of $\geq$2 appliances in all cases, our model predicts an average of 1 appliance per household for all but AC units by 2025. This is true of refrigerators, and washing machines, which already near an average of 1, as well as water heaters, which were close behind. It is also true of computers and microwaves, which had only achieved about 70% penetration by 2009 but were on a clear upward trajectory. Going back to the underlying data, we note that there are very few households with more than one of these appliances in the sample, even at high-income levels, reinforcing the basis for “1” as the limiting value. Meanwhile, air conditioners remain on an upward trend and approach 1.8 ACs per household at the end of the forecast. Chinese apartments usually have separate AC units for bedrooms, not a central AC system as is typical in US homes. Central AC systems may become more popular among new buildings in the future, but this could not be analyzed in our framework as we do not observe whether a household has a central AC system in the data.

4.3. Electricity consumption results

Table 4 reports the OLS estimation results for the electricity consumption model. As in the earlier models, all of our electricity models include linear and quadratic income terms, allowing the income elasticity to vary with income level. Model 1 includes only the electricity price, income, year fixed effects and regional indicators as explanatory variables to serve as a “simple model” benchmark. Model 2 includes the variables in Model 1, plus dwelling size and appliance stocks, which in addition to income, are the necessary factors driving subsequent electricity forecasts. Models 3, 4 and 5 are the main

Table 4
OLS estimation results for electricity consumption (dependent variable = ln(electricity consumption)).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
lnp	−0.720***	−0.655***	−0.659***	−0.655***	−0.652***	−0.610***
	(0.120)	(0.136)	(0.113)	(0.113)	(0.112)	(0.0946)
lny_c	0.497***	0.330***	0.322***	0.323***	0.322***	0.337***
	(0.0173)	(0.0143)	(0.0161)	(0.0161)	(0.0161)	(0.0157)
lny_c^2	−0.0884***	−0.0877***	−0.0899***	−0.0900***	−0.0901***	−0.0917***
	(0.0102)	(0.00890)	(0.00893)	(0.00892)	(0.00893)	(0.00986)
lndwelling_sz_c		0.0739**	0.0657**	0.0657**	0.0658**	0.0652**
		(0.0265)	(0.0235)	(0.0235)	(0.0234)	(0.0210)
fridge		0.191***	0.187***	0.187***	0.187***	0.192***
		(0.0187)	(0.0194)	(0.0194)	(0.0195)	(0.0175)
comp		0.0469***	0.0589***	0.0589***	0.0589***	0.0678***
		(0.00740)	(0.00704)	(0.00707)	(0.00708)	(0.00744)
mwave		0.0729***	0.0838***	0.0839***	0.0843***	0.0902***
		(0.0132)	(0.0120)	(0.0120)	(0.0120)	(0.0123)
wheater		0.0407**	0.0506***	0.0506***	0.0506***	0.0507***
		(0.0147)	(0.0136)	(0.0136)	(0.0136)	(0.0132)
washer		0.0358**	0.0319*	0.0323*	0.0325*	0.0160
		(0.0130)	(0.0125)	(0.0126)	(0.0126)	(0.0122)
ac		0.109***	0.0744***	0.0734***	0.0733***	0.0765***
		(0.00947)	(0.0167)	(0.0164)	(0.0165)	(0.0163)
household_sz			0.0569***	0.0568***	0.0568***	0.0575***
			(0.00510)	(0.00511)	(0.00513)	(0.00490)
age_over50			0.0868***	0.0869***	0.0871***	0.0908***
			(0.00977)	(0.00973)	(0.00973)	(0.0118)
edu_mid			−0.00978	−0.00923	−0.00894	−0.00496
			(0.0137)	(0.0137)	(0.0137)	(0.0153)
edu_high			−0.0212	−0.0202	−0.0195	−0.0205
			(0.0139)	(0.0139)	(0.0139)	(0.0186)
edu_coll			−0.0565***	−0.0557***	−0.0549***	−0.0599**
			(0.0156)	(0.0156)	(0.0156)	(0.0198)
hdd (in 1000)			0.0134	0.00981	0.0103	−0.0322
			(0.0259)	(0.0237)	(0.0235)	(0.0301)
cdd (in 1000)			0.193	0.187	0.193	0.415**
			(0.153)	(0.137)	(0.135)	(0.156)
hdd_heat_ac			0.0449**	0.0446**	0.0444**	0.0789***
			(0.0147)	(0.0146)	(0.0146)	(0.0226)
hdd_heat_gaswater			−0.0224	−0.0219	−0.0218	0.0689
			(0.0312)	(0.0312)	(0.0311)	(0.0385)
hdd_heat_other			−0.0558	−0.0546	−0.0541	0.00296
			(0.0341)	(0.0342)	(0.0343)	(0.0359)
cdd_ac			0.0865*	0.0883*	0.0890*	0.0895*
			(0.0424)	(0.0420)	(0.0423)	(0.0439)
b_1 *(time trend)*				0.541*	0.0342***	
				(0.253)	(0.00396)	
b_2				0.0836		
				(0.0513)		
α	6.283***	5.688***	5.650***	5.966***	5.433***	5.883***
	(0.103)	(0.155)	(0.188)	(0.292)	(0.184)	(0.140)
heating types	*NO*	*NO*	*YES*	*YES*	*YES*	*YES*
year dummies	*YES*	*YES*	*YES*	*NO*	*NO*	*YES*
regional dummies	*YES*	*YES*	*YES*	*YES*	*YES*	*NO*
N	122252	122252	122252	122252	122252	122252
adj. R^2	0.342	0.385	0.397	0.397	0.397	0.387

Clustered standard errors in parentheses.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

specifications, additionally controlling for demographic variables, as well as heating type, weather variables, and interactions among them. Regional indicators are included in all specifications.

Another key difference is what we assume about the time trend. As with dwelling size, in Model 3 we consider year fixed effects (which provide no independent time trend), while Models 4 and 5 include non-linear and linear trends respectively. Model 6 excludes regional indicators, allowing us to examine how parameters related to weather are affected by requiring them to predict between as well as within region variations.

Table 5
Residential electricity consumption elasticities from previous studies.

Sources	Price elasticity	Income elasticity	Income ($2010)*	Period	Country
Cao et al. (2016)	−0.55	0.65	1115	2002–2009	China
Zhou and Teng (2013)	−0.35	0.14	1333	2007–2009	China
Shi et al. (2012), small sample	−2.48	0.06	1385	2008–2009	China
Our estimate for year 2009	−0.65	0.28	1444	2009	China
Our estimate for year 2025	−0.65	0.11	3838	2025 (E)	China
Baker et al. (1989)	−0.76	0.13	11,145	1972–1983	UK
Dubin and McFadden (1984)	−0.26	0.02	16,027	1975	US
Parti and Parti (1980)	−0.58	0.15	16,027	1975	US
Hirst et al. (1982)	−0.67	0.16	17,974	1978–1979	US
Bernard et al. (1996), IV	−0.67	0.14	18,264	1986–1989	Canada
Nesbakken (1999), short run	−0.45	0.01	23,580	1993–1995	Norway
Fell et al. (2014)	−0.48	0.01	33,600	2006–2008	US

Note: Income (in 2010 dollar) is the average of the final consumption per capita for the study period based on the World Bank data. For our estimate in 2025, the income is calculated by applying our growth rate in consumption per capita to the 2009 income level.
Link: http://data.worldbank.org/indicator/NE.CON.PRVT.PC.KD.

4.3.1. *Price and income elasticities*

With the exception of Model 1, the price and income elasticities are relatively consistent across all models, with a price elasticity of −0.66 and income elasticity of 0.32 at the sample average income (recall the income variable has been normalized to sample mean, so the log is zero at that point). As average income rises over the sample period and forecast horizon, however, this income elasticity falls based on the quadratic term. It is 0.28 for 2009 and falls to 0.11 by 2025. These elasticities could be viewed as "short-term" responses at the intensive margin, as dwelling size, major appliances and other demographic characteristics are controlled in the specifications. Model 1 does not control for these features, so the price elasticity of −0.72 and income elasticity of 0.50 (at the sample average) account for both higher utilization of existing appliances (intensive margin) and higher adoption rate of major appliances and larger houses (extensive margin). These might be viewed as "long-term" elasticities. As figuring out how much dwelling size and appliances adoption contribute to demand growth is critical to policy design, Models 3–5 are more informative in this sense and are therefore used in the forecast. If we use city-level average electricity price (not reported in the table) the price elasticities are slightly larger, ranging from −0.64 to −0.86.

To gain a better sense of how these estimates compare with previous work, we summarize in Table 5 some relevant previous studies on price and income elasticities in the residential sector from China and other countries. We restrict our selection to ten papers that use micro-level data or papers having elasticity estimates for residential electricity in China with comparable methods. Estimates of price elasticity range from −2.5 to −0.26; if we set aside the two lowest and two highest estimates, the range is a much tighter −0.45 to −0.67, putting our estimates at the upper end of that range.

Given the sensitivity of income elasticity to income level, we use Fig. 5 to show the values from columns 2 and 3 of Table 5 in a figure alongside our estimates for 2009 and 2025. Here, the average income for the country and time period is calculated based on the final household consumption per capita data from the World Bank to make the income elasticities more comparable. Estimates of income elasticity are generally below 0.3, which is consistent with the fact that electricity consumption is typically a smaller share of richer household's expenditures. More importantly, we can see that income elasticities eventually fall to zero as income rises, usually at an income level higher than $15,000 per capita in most studies.

Estimated future levels of income elasticity for Chinese households based on projected future income levels are consistent with the income elasticity pattern across countries of various income level. Our estimates predict that the elasticity will approach zero when income reaches $8200 or 57,700 RMB per capita (in 2010 RMB) in urban China. Given the income growth rate we use in the forecast, this will be achieved around the year 2040. Income growth could still drive electricity growth through larger dwellings and more appliances beyond this point.[21]

Focusing on a more specific comparison, Zhou and Teng (2013) use data on 2007–2009 in Sichuan Province that is drawn from the same dataset as ours. For comparison purpose, we can apply our quadratic estimates to the mean real income for 2007–2009 (RMB 26609), yielding an income elasticity of 0.29. They find a significantly lower income elasticity, about 0.14. One reason for the discrepancy may be the absence of a quadratic income term in their specification. As they add demographic variables, their income elasticity falls. Intuitively, education indicators could be picking up what would otherwise be non-linear income effects—particularly, the higher income elasticity at middle to low-income levels. Their income elasticity, therefore, is more heavily weighted towards high-income levels (which is lower), where education does not vary. Cao et al. (2016) also draw from the same data. They report a price elasticity of −0.55 and an income elasticity of 0.65 for the middle-income group using the same data set. Their income elasticity is roughly double ours over the historic sample; the difference may be due to their separation into income groups and use of a two-stage budgeting AIDS model which differs substantially

[21] If one employs the income elasticity from Model 1, which does not control directly for dwelling size and appliance stocks, the income elasticity reaches zero by 2050.

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

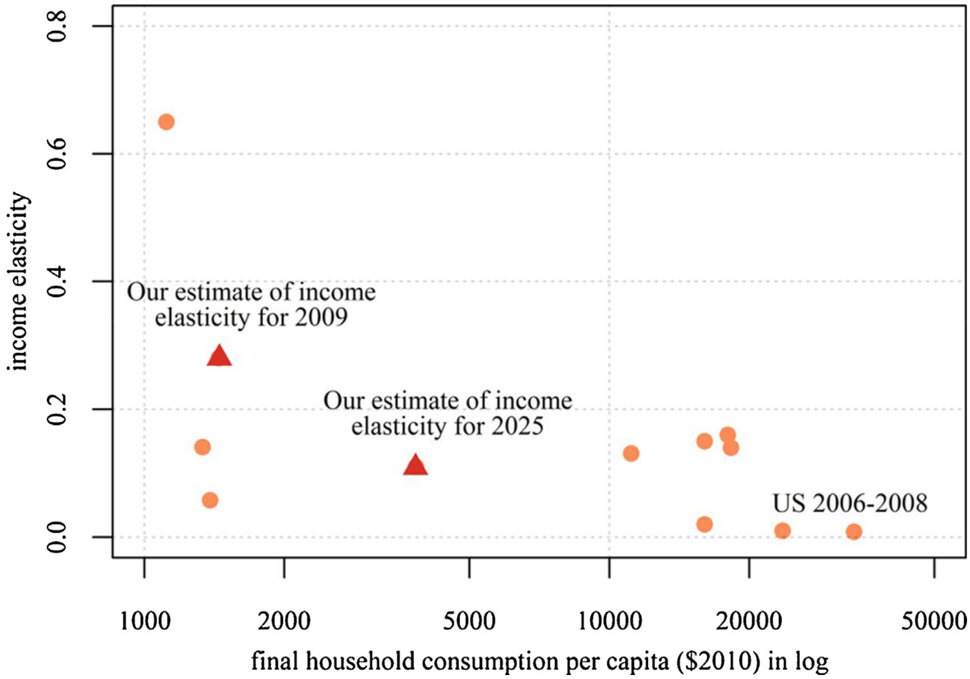

Fig. 5. Income elasticities versus income levels from comparable studies.

from our reduced form model. Finally, a recent paper by Du et al. (2015) estimates the elasticities by tier pricing structure and provides an income elasticity of roughly 0.10–0.15, and a price elasticity of roughly −0.65 to −1.3. The set of results are comparable in scale to ours especially given that they study more recent years, although using a less representative sample.

4.3.2. Appliances (except AC) and electricity consumption

An important advantage of our modeling effort is the ability to explicitly model the contribution of appliance adoption to electricity consumption. We find that all appliances contribute positively to electricity demand and the coefficients are all highly statistically significant except for washing machines. The coefficients are relatively constant across the three different time trend models (Models 3–5), but all appliances except refrigerators have some variation with the removal of demographic controls (Model 2) or regional controls (Model 6). This is not surprising as some appliances, like computers, are highly correlated with education.

Among the non-AC appliances, refrigerators consume the highest level of electricity, raising consumption by 19%. Given the saturation of refrigerator ownership, this will have little impact on our projections of ownership, but the saturation does point to the potential savings if the efficiency of refrigerators can be improved. For example, refrigerator efficiency doubled in the United States from 1987 to 2012 (Mauer et al., 2013). If that same improvement occurred in China over our forecast period, it could reduce Chinese residential electricity usage by 10%; the value of these reductions would need to be compared to the costs to assess the net benefits of such efficiency improvements. Other appliances increase electricity consumption by much smaller amounts, from 3% to 8%.

4.3.3. AC, weather and heating types and electricity consumption

In the preferred models (3–5), the coefficient on AC units is about 7%. However, this is not the whole story, as the interaction with heating and cooling degree days adds another 7%. That is, we find statistically significant effects of AC unit adoption interacted with cooling degree days and with heating degree days, where the latter is further interacted with whether the household uses its AC unit for heating. Here, we note that households in our sample report whether they heat with (a) their AC unit, (b) gas or hot water, (c) some other fuel, or (d) nothing, with the last serving as our reference.

To compute the effect of these interactions, we consider the estimated coefficient multiplied by the average values of the variables interacted with AC. For example, there is an average of 280 cooling degree days in the sample, so the first interaction is $0.089 \times 0.28 = 2.5\%$. Meanwhile, there are 3700 average heating degree days, and 0.27 of households use AC for heating, so the second interaction is $0.045 \times 0.27 \times 3.7 = 4.5\%$. (These numbers are also confirmed by simulation.) Of course, it is not clear whether future AC adopters will be above or below the current mean values for heating and cooling degree days of current AC owners.

In addition to the estimated effect of AC adoption being about 14%, it is important to recognize from the preceding discussion that AC units are the one appliance expected to be still growing at the end of the forecast horizon. While microwaves and computers see adoption rates increase by around 30% from 2009 to 2025, the adoption then flattens out. The adoption

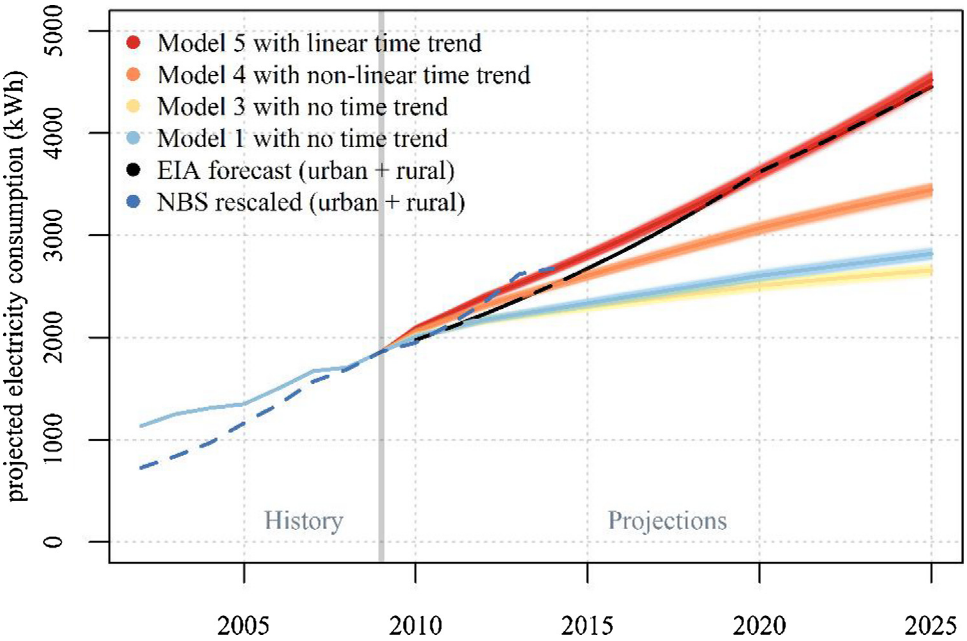

Fig. 6. Mean electricity consumption (in kWh) projected from 2009 using alternative time trends. Shaded regions indicate variation based on uncertain demographics.

of AC units, in contrast, increases about 50% over the period and is projected to continue to grow at almost the same rate. Assuming the average number of units per household rises by one every 20–25 years, that would be a further 14% increase in electricity usage by 2050. Thus, AC units could be another focus for energy efficiency policies, though improvements in the US over the past 20 years have been smaller for AC than for refrigerators (Mauer et al., 2013).

Finally, it is worth noting that inclusion of demographic, weather, and regional controls matter for these estimates. The simpler model without weather and demographics suggests an AC elasticity closer to 11%, versus the 14% above. On the other hand, removing the regional controls (and introducing considerably more variation in degree days), raises the coefficient on heating-AC interaction from 0.045 to 0.079, suggesting and effect closer to 17% (the change, 0.034, $\times$ 0.27 $\times$ 3.7, from above, equals +3%). Generally, the other weather-related variables (beyond those interacted with AC) are not statistically significant. This is perhaps not surprising as it is unclear how they would relate to electricity use, independent of space conditioning.

4.3.4. Demographic features

Dwelling size has a positive but small influence on electricity use. The coefficient is roughly 0.07 in all models. Given the increase in dwelling size over the forecast horizon of roughly 60% with the most optimistic trend assumption, this would contribute only 3% to electricity demand over the forecast horizon. Household size and an indicator of whether the head of household is over 50 contribute positively to electricity consumption. One more person in the household increases the electricity consumption by 5%. However, on a *per person* basis, larger households save energy as more appliances, for example, refrigerators and washing machines, can be shared. These results are similar to Zhou and Teng (2013).

We also see that higher education levels lead to lower electricity use—almost 6% for college-educated heads of households. This is consistent with the idea that more educated households may have increased awareness of electricity conservation methods or a preference for more efficient appliances. While this contrasts with Zhou and Teng (2013), as noted above, the absence of a quadratic income variable likely implies a bias in that study. Educational variables, which are highly correlated with income, may be picking up non-linear income effects rather than the independent effect of education holding income constant.

5. Forecast electricity consumption and discussion

An important goal of this study is to generate forecasts of average urban household electricity demand. We also seek to understand the components of existing demand and future growth, and the potential implications of policy interventions or electricity market reforms. Fig. 6 depicts the forecasted average urban household's electricity consumption based on Models

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003

Table 6
Forecast sensitivity to non-income changes.

Variable name	Assumed change in the variable 2010–2025	Corresponding projected change in electricity consumption 2010–2025
Household size	−13%	−2.2%
Age of household head >50	16%	1.4%
Education:		
Primary or below	2%	
Middle school or equivalent	−3%	0.03%
High school or equivalent	−11%	0.22%
College or above	12%	−0.67%
Heating types:		
None	−42%	
Heating by AC	47%	1.3%
Heating by gas or water	3%	−0.4%
Heating by other	−8%	−0.1%
Heating degree days	46%	1.6%
Cooling degree days	12%	0.7%
		Total: 2.1%
Income	165%	27%

Note: Income is assumed to follow the projected per capita income growth in IEA's World Energy Model 2015. The assumed changes in other variables follow the historical trend observed in the data. We use the coefficients from Table 4 (Model 4) to calculate the corresponding projected change.

3–5, as well as Model 1, a recent EIA forecast (EIA, 2016) and the reported data by NBS[22] (rescaled to match our data in 2009). Dwelling size is forecast using Model B (the non-linear model) described in Section 4.1, appliances as described in Section 4.2, and income based on the WEO sources as described at the beginning of Section 4. All other variables are assumed to maintain at the same level of those in the year 2009. Hence, income drives electricity growth both directly and indirectly through increase in dwelling size and appliances, permitting a decomposition described below.

We first note the importance of time trend assumptions. Despite efforts to build a model that accommodates time-varying determinants of electricity demand growth—income, dwelling size, and appliance stocks—time trend assumptions still make a significant difference to the forecast results. The model with year fixed effects ignores any autonomous trend and predicts a 43% increase in household electricity consumption from 2010 to 2025.[23] The growth rate is much higher using the models that instead estimate and project such trends. Electricity consumption in 2025 increases by 143% over 2009 using a linear time trend model that simply extrapolates the average 2002–2009 trend. Given we observe a slight slowdown in growth in the historical data, we also estimate a non-linear trend model. This model projects 85% growth in electricity consumption. It is worth noting that our linear trend model produces results relatively close to the EIA projection.[24] This, of course, does not mean that the linear model is superior, but simply implies that the forecasts match. Ultimately, our best guess is somewhere between Models 4 and 5. It is unrealistic to assume the observed autonomous trend in the data stops in 2009. However, it is not clear whether we should expect the trend to slow or not over the forecast horizon.

In Table 6, we summarize the forecast sensitivity to non-income time trends using coefficient estimates from electricity Model 4. Rather than fixing demographic variables at their 2009 levels, we instead consider the historical trend observed in our data from 2002 to 2009. Extrapolating those trends, we estimate the likely demographic change from 2010 to 2025, reported in the first column of Table 6. We then multiply by the corresponding coefficient in Model 4 to produce a projected effect on the electricity demand projection. All of these demographic trends lead to very limited changes in electricity consumption (−2.3% to 1.7%).

We considered other approaches to try to reduce uncertainty about the trend model. We compared our projections for 2010–2014 to more recently available national totals published by the NBS. However, these include both the urban and rural residents in all provinces, thus do not match perfectly to our scope. We also considered cross-validation using our sample, but the relatively short time series of 9 years makes it difficult to distinguish the linear and non-linear models.[25] Thus we ultimately prefer to highlight both the non-linear and linear time trend models as our preferred models, with projections

[22] The total residential electricity consumption is from Tables 4–13 Consumption of Electricity and Its Main Varieties by Sector, China Energy Statistical Yearbook 2015. The per household consumption is calculated using the population and household size from the online database on NBS website.

[23] Based on the official data from the NBS, consumption expenditures grew by 10.8% annually from 2010 to 2012 and by 9.4% annually from 2012 to 2014, compared to the 9.5% and 6% forecasted. Putting this actual growth rate would increase the forecasted growth rate (2010–2025) from 43% to 47%. Urban household disposable income grew by 12.7% annually from 2010 to 2012; using this actual growth would increase the forecasted growth rate from 43% to 45%. Per capita income grew even faster, and using it would increase the forecasted growth rate from 43% to 48%.

[24] EIA forecasts the population and residential electricity to the year 2040. We first calculate the annual growth rate implied by the calculated per capita residential electricity and then apply the annual growth rate from 2010 to our historical data in the year 2009 to make the forecasts more comparable to our results. However, it is worth noting that EIA's forecasts do not differentiate between urban and rural.

[25] We tried cross-validation by estimating the models with the 2002–2005 sample and forecast to 2009 using the same income assumption from IEA and holding other variables fixed at 2005 level. Linear-trend model performs relatively well. However, this does not give us the confidence to choose the

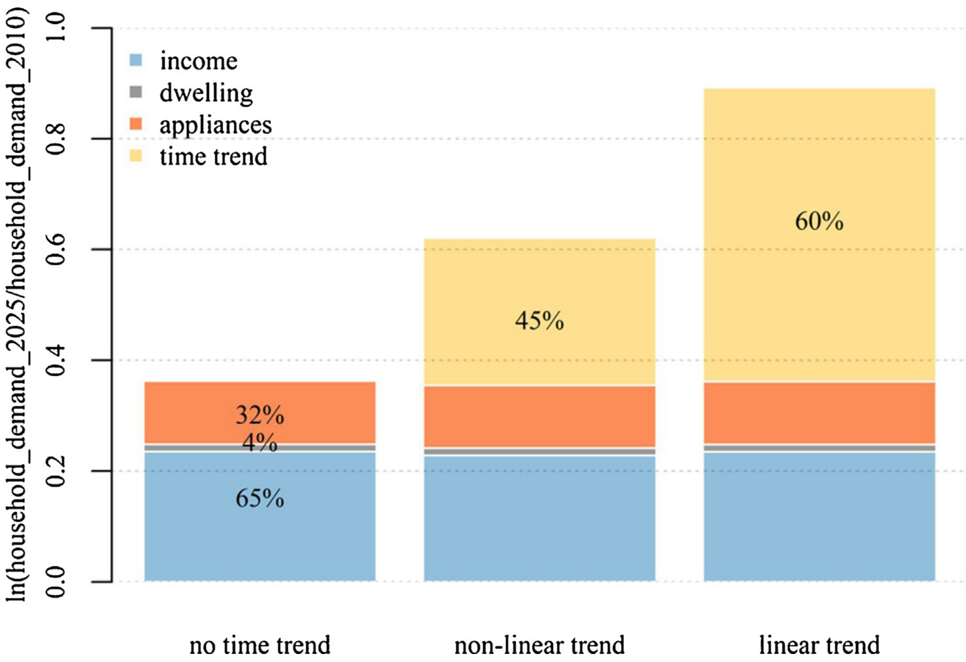

Fig. 7. Decomposition of projected electricity growth for 2010–2025 (percentages of each model total).

of between 85% and 143% per capita growth from 2009 to 2025. It is feasible, through our model, to highlight the important assumptions that have to be made, including the time trend decision discussed above.

One question that motivated our work was whether creating a more detailed and disaggregated model of energy demand would improve demand forecasts. As we can see in Fig. 6, there is little difference between using Model 1 and Model 3 as forecasting tools. Recall that the main difference is that Model 3 controls appliances and other socioeconomic characteristics while Model 1 only control price and income. This suggests that the quadratic income specification in the simple model is capable of capturing most of the total contribution of income to electricity consumption growth, directly and indirectly. In other words, controlling the dwelling size, appliances and other demographic variables in a more complicated model does not improve the forecasts substantially—so long as dwelling size and appliance adoption are similarly forecasted using income. It is still important to separate income from time trends. The income elasticity halves from 2009 to 2025, regardless of the autonomous time trend. But as household-level data are not available for most researchers and policy makers, this result suggests that aggregated data at city-level or province-level might be sufficient. That is, such aggregate data might resolve the more singular distinction between income effects and autonomous trend and provide a basis for simple demand forecasting. However, micro-level data are still necessary to understand the relative importance of other components of growth and the potential for policy impacts, as discussed next.

5.1. Decomposition of factors driving electricity consumption growth

Fig. 7 addresses the question of how much the various components of growth contribute to the total. This figure shows the relative contribution to growth over the period 2009–2025 from four factors: appliance adoption, dwelling size, income effects (conditional on adoption and dwelling size), and unrelated time trends. The first two are extensive margins that increase the size of homes and the stock of major electricity-using devices, while the third includes more intensive use of the same stock. It is unclear whether autonomous time trends are extensive or intensive.

The figure suggests that appliance adoption contributes as much as 30% of the increase in energy demand, while dwelling size contributes at most 6%. Meanwhile, the contribution of an autonomous time trend ranges from zero (first column for year fixed effects) to more than equal to the combined model components of income, appliances, and dwelling size (last column). Despite the previous point that the aggregate growth projections are similar using income alone, we can now see the pathways driving the projected growth.

From earlier discussions, we know that the extensive appliance contribution is largely AC units. While refrigerators may be an obvious choice for energy efficiency policy given their larger total contribution and known potential, AC may be particularly important given the significant addition of units over the next 15 years, as well as the potential for a shift to

linear-trend model as estimating a non-linear trend with four-year data is challenging. It is possible that the non-linear trend estimated from the eight-year data will perform better in the longer horizon.

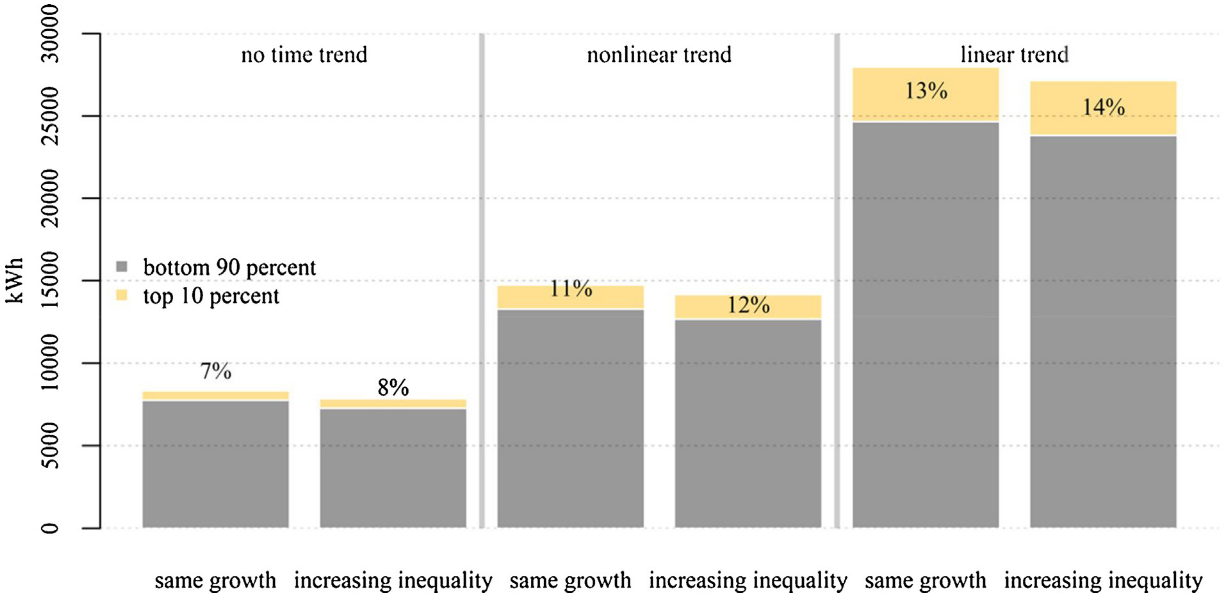

Fig. 8. Projected electricity growth (in kWh) by income groups for 2010–2025.

central cooling technologies. Not only is there an intensive margin, as reflected by the parameter estimates discussed in Section 4.3.3, the detailed projections reveal an important extensive margin.

5.2. Income distribution

A final question is the possible effect of the *distribution* of economic growth. Auffhammer and Wolfram (2014) illustrate that different shapes of the underlying distribution of income could lead to different demand estimates, even if the mean income is the same. This would be true in our model, where we allow the income elasticity to vary with income. Given we observe changing income patterns in our data from 2002 to 2009, we have some basis for considering how income growth might vary across income levels in the forecast period.

For example, we could assume the projected 165% increase in per capita income growth over 2010–2025 applies equally to all groups, keeping the relative distribution fixed as it was in 2009. Or, we could assume a higher growth rate for wealthier households. This is consistent with recent experience in the United States (Heathcote et al., 2010), the literature on income inequality in China (Wu and Perloff, 2005) and our data. That is, over 2002–2009, we see 30% higher growth in the top decile of our data compared to the lower nine deciles. Consider what happens if, instead of assuming the same growth for all deciles, we assume the top decile grows 30% faster than the lower nine deciles over 2010–2025. We maintain the same aggregate growth of 165% over that time period.

Fig. 8 shows the results. For each of the three trend assumptions, a pair of bars shows the difference in electricity demand growth between assuming (a) everyone's income grows 165%, versus (b) the top decile grows 220% while the bottom grows 145%. Each bar also indicates the contribution of the top 10% (top, labeled) versus the bottom 90% (bottom).

In general, we see a very small effect from differing growth assumptions within each pair of bars for the different trend assumptions. Assuming more growth in the top decile clearly lowers demand growth. Every dollar in income growth shifted from the poor to the rich means (a) a larger percent decrease for the poor along with a smaller percent increase for the rich, and (b) that larger percent decrease for the poor is applied to a larger energy elasticity for lower-income households. What was unclear prior to the simulation was that the effect would be small given the differential growth rates we used. Moving from left to right, the autonomous time trend is assumed to apply equally to all income groups and therefore dilutes whatever effect we see in the left two bars. Note a more extreme assumption (e.g., no income growth among the poor) would presumably show more marked effects.

6. Conclusions and future efforts

Forecasting electricity demand and the potential impact of alternative policies will be an important task as China grapples with both domestic development needs and increasing environmental concerns at home and abroad. Growth in residential electricity use is projected to contribute roughly one-third of total demand growth through 2025, yet is relatively unstudied. This paper has provided the first attempt to use household-level data to construct such forecasts. We also use that data to decompose future electricity forecasts into contributions from appliance adoption, dwelling size increases, direct income effects, and potential autonomous time trends.

Autonomous time trends—that is, growth that we cannot link to income or other observables in the data—creates considerable uncertainty, with the potential to more than double growth from estimable sources. These autonomous trends could be due to the changing relative prices of housing and appliances, changing quality of appliances—e.g., larger refrigerators and washing machines or more powerful microwave ovens—or due to adopting more energy-using habits at higher income levels. Given we think it is unrealistic to assume such autonomous trends are zero, our preferred range of projected per capita demand growth is between 85% and 143% over 2009–2025 using alternative linear and non-linear autonomous trends.

Consistent with other work, we find electricity price elasticities of −0.7. A price increase of 15%, consistent with a $10 per ton price on carbon dioxide emissions, would lower demand by about 10%. This would be one possible policy, along with raising consumer prices through market liberalization in the power sector, to reduce demand growth.

Perhaps most interestingly, appliance adoption—particularly a larger number of AC units—drives up to a third of future electricity growth. Efficiency policies targeted at AC units would have a high potential impact as so many new ones are expected to drive energy demand. Meanwhile, refrigerators appear to be a large source of current household energy use. Refrigerator ownership raises household electricity demand by an estimated 19%. Thus reducing energy use of these appliances in half would reduce electricity demand by 10%. However, refrigerators have also saturated the market. Nearly 100% of households have one. Few are seeking to have two or more.

These back-of-the-envelope calculations of potential policy impacts could be expanded in future analysis. Such an analysis could consider more explicitly the turnover of appliances and the effect of energy and product prices, as well as electricity generation costs and environmental impacts. This would allow a better comparison of price changes, efficiency policies, or other policies to reduce electricity use. One could also focus more specifically on the role of residential energy tariffs and proposed reforms. Finally, one could consider how changing temperatures due to climate change would exacerbate the trend towards more AC adoption and higher energy use.

Looking forward, it would also be valuable to increase understanding about the residual time trend in household electricity use. Is there a difference in the income elasticity observed between households at a point in time (which we estimate) and those observed within a household over time (which would require a panel data set)? It could also be useful to understand whether and how elasticities vary through quantile regression or further sub-group analyses. These and other questions present exciting areas for future work.

Appendix A. Household data cleaning process

The cross-sectional Chinese Urban Household Survey (CUHS) data have two different data sets: one dataset on household features (e.g., household size, dwelling size, electricity consumption) and another dataset on household head characteristics (e.g., age and education level of household head). We first merged the two datasets by household identity generated from household code and regional code: 510420 out of 510474 observations are merged successfully. Note that it is possible that two or more people are matched to one household because multiple people are surveyed in the household. The people merged to one household would share the same household-level characteristics while differ in personal characteristics. We calculate the number of old people and the number of young people matched to each household. This information is added to the dataset as a household-level feature for each observation.

The cleaning process for the merged dataset is shown in the following table. The steps are conducted sequentially. For each step, we present the number of the observations dropped in the step and the share of the dropped observations among the total 510420 successfully merged observations. Note that we only focus on the demographic information of the household head because they are assumed to be the decision maker of the household. The observations before the cleaning process is on individual level while the unit of observations after the cleaning process is the household level.

	Reason of dropping	Number dropped	Share(%)
Step 1.	No demographic information	125310	14.55
Step 2.	Zero electricity expenditure	11896	2.33
Step 3.	Zero household size	2569	0.50
Step 4.	Zero dwelling size	182	0.04
Step 5.	Age less than 16	50754	9.94
Step 6.	Dwelling size larger than 300 square meters	1870	0.37
Step 7.	More than 6 old people in the household	4296	0.84
Step 8.	Heat by AC but do not own AC	1093	0.21
Step 9.	Have missing values for the variables	54	0.01
Step 10.	Not household head	190144	37.24
	Total number left:	122252	23.95

References

Alberini, A., Filippini, M., 2011. Response of residential electricity demand to price: The effect of measurement error. Energy Econ. 33 (5), 889–895.
Auffhammer, M., Wolfram, C.D., 2014. Powering up China: income distributions and residential electricity consumption. Am. Econ. Rev. 104 (5), 575–580.
Paul Baker, Paul, Blundell, Richard, Micklewright, John, 1989. Modelling household energy expenditures using Micro-data. The Economic Journal 99 (397), 720–738.

Bernard, J.T., Bolduc, D., Belanger, D., 1996. Quebec residential electricity demand: a microeconometric approach. Canad. J. Econ., 92–113.
Brandt, L., Holz, C.A., 2006. Spatial price differences in China: estimates and implications. Econ. Dev. Cult. Change 55 (1), 43–86.
Cao, J., Ho, M., Liang, H., 2016. Household Energy Demand in Urban China : accounting for regional prices and rapid income change. Energy J. 37, http://dx.doi.org/10.5547/01956574.37.SI1.jcao.
Du, G., Lin, W., Sun, C., Zhang, D., 2015. Residential electricity consumption after the reform of tiered pricing for household electricity in China. Appl. Energy 157, 276–283.
Dubin, J.A., McFadden, D.L., 1984. An econometric analysis of residential electric appliance holdings and consumption. Econometrica: J. Econ. Soc., 345–362.
Energy Information Administration (EIA), 2015. International Energy and Data Analysis (China). http://www.eia.gov/beta/international/analysis_includes/countries_long/China/china.pdf.
Energy Information Administration (EIA), 2016. International Energy Outlook. https://www.eia.gov/outlooks/ieo/pdf/0484(2016).pdf.
Espey, J.A., Espey, M., 2004. Turning on the lights: A meta-analysis of residential electricity demand elasticities. J. Agric. Appl. Econ. 36 (1), 65–81.
Fell, H., Li, S., Paul, A., 2014. A new look at residential electricity demand using household expenditure data. Int. J. Ind Organiz 33, 37–47.
Fisher, F.M., Kaysen, C., 1962. A Study in Econometrics: The Demand for Electricity in the United States. North-Holland Publishing Co, Amsterdam.
Halvorsen, R., 1975. Residential demand for electric energy. Rev. Econ. Stat., 12–18.
Heathcote, J., Perri, F., Violante, G.L., 2010. Unequal we stand: an empirical analysis of economic inequality in the United States, 1967–2006. Rev. Econ. Dyn. 13 (1), 15–51.
Hirst, Eric, Goeltz, R., Carney, J., 1982. Residential Energy Use: Analysis of Disaggregate Data.
International Energy Agency (IEA), 2014. World Energy Outlook.
Jones, R.V., Lomas, K.J., 2015. Determinants of high electrical energy demand in UK homes: socio-economic and dwelling characteristics. Energy Build. 101, 24–34.
Larsen, B.M., Nesbakken, R., 2004. Household electricity end-use consumption: results from econometric and engineering models. Energy Econ. 26 (2), 179–200, http://dx.doi.org/10.1016/j.eneco.2004.02.001.
Leahy, E., Lyons, S., 2010. Energy use and appliance ownership in Ireland. Energy Policy 38 (8), 4265–4279.
Mauer, Joanna, deLaski, Andrew, Nadel, Steven, Fryer, Anthony, Young, Rachel, 2013. Better Appliances: An Analysis of Performance, Features, and Price as Efficiency Has Improved. ACEEE, Washington (ACEEE Research Report A132).
McNeil, M.A., Letschert, V.E., 2010. Modeling diffusion of electrical appliances in the residential sector. Energy Build. 42 (6), 783–790.
Nesbakken, R., 1999. Price sensitivity of residential energy consumption in Norway. Energy Econ. 21, 493–515.
O'neill, B.C., Chen, B.S., 2002. Demographic determinants of household energy use in the United States. Popul. Dev. Rev. 28, 53–88.
Parti, M., Parti, C., 1980. The total and appliance-specific conditional demand for electricity in the household sector. Bell J. Econ., 309–321.
Reiss, P.C., 2005. Household electricity demand revisited. Rev. Econ. Stud. 72, 853–883.
Shi, G., Zheng, X., Song, F., 2012. Estimating elasticity for residential electricity demand in China. Sci. World J. 2012.
Shi, Y., Gao, X., Xu, Y., Giorgi, F., Chen, D., 2016. Effects of climate change on heating and cooling degree days and potential energy demand in the household sector of China. Climate Res. 67 (2), 135–149.
Swan, L.G., Ugursal, V.I., 2009. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. Renew. Sustain. Energy Rev. 13 (8), 1819–1835, http://dx.doi.org/10.1016/j.rser.2008.09.033.
Taylor, L., Blattenberg, G.R., Rennhack, R., 1984. Residential demand for energy in the United States: empirical results for electricity. In: Moroney, J.R. (Ed.), Advances in the Economics of Energy and Resources, vol 5. JAI Press, Greenwich, CT, pp. 85–102.
Wu, X., Perloff, J.M., 2005. China's income distribution, 1985–2001. Rev. Econ. Stat. 87 (4), 763–775.
Zhou, S., Teng, F., 2013. Estimation of urban residential electricity demand in China using household survey data. Energy Policy 61, 394–402, http://dx.doi.org/10.1016/j.enpol.2013.06.092.

Please cite this article in press as: Cao, J., et al., Chinese residential electricity consumption: Estimation and forecast using micro-data. Resource Energy Econ. (2017), https://doi.org/10.1016/j.reseneeco.2017.10.003