Effects of building size growth on residential energy efficiency and conservation in California

Eric D. Fournier a, Felicia Federico a, Erik Porse a,b, Stephanie Pincetl a,n

a Institute of the Environment and Sustainability, University of California Los Angeles, La Kretz Hall, Suite 300, 619 Charles E. Young Drive E., Los Angeles, CA 90095-1496, United States
b Office of Water Programs, California State University Sacramento, 6000 J Street, Sacramento, CA 95819-6025, United States

HIGHLIGHTS

- EUI improvements are not keeping pace with the size growth of new homes in LAC.
- LAC neighborhood home EUI trends strongly correlate with vintage and income levels.
- Wealthy area homes are proportionally larger than they are more energy efficient.
- Limiting size growth among new homes could result in substantial energy savings.

ARTICLE INFO

Keywords:
Residential buildings
Energy efficiency
Energy conservation
Electricity
Natural gas
Account level data

ABSTRACT

Many utility and government programs exist to promote energy efficiency (EE) in residential buildings. While programs have succeeded in reducing per square foot energy usage intensity (EUI), they do not necessarily promote conservation, in terms of reduced total energy consumption. Using statistical analysis and data mining techniques, we examined relationships between home size, electricity and natural gas EUIs, and neighborhood level socio-economic attributes among ∼1.3 million single-family homes in Los Angeles County (LAC). We observed that among homes constructed between 1900 and 2010, the growth in median home size by construction vintage year has outpaced combined EUI reductions by 60%. Results of a Monte-Carlo sampling procedure derived from these observed trends indicate that past historical energy savings within LAC, attributable to state mandated EE policies, could have been equivalently achieved by constraining growth in the size of newly constructed homes. These findings have significant implications for the design of future energy conservation policies within growing urban areas.

1. Introduction

Each year within the United States, billions of dollars in energy utility ratepayer funds are spent on demand side management programs focused on improving building energy efficiency (EE) or, equivalently, on reducing per square foot energy usage intensity (EUI) [1]. This pursuit of building energy conservation through these marginal EE gains is attractive in theory because it does not require building occupants to reduce their effective demand for energy services in order to achieve reductions in associated greenhouse gas (GHG) emissions. However, as numerous studies have shown, improvements in EE do not guarantee increased levels of energy conservation - a phenomenon commonly referred to within the literature as energy rebound or Jevon’s Paradox [2-9].

This study investigated differential changes in EUI and average built square footage, relative to construction vintage, among a large sample of single-family homes located within Los Angeles County (LAC). This investigation was based upon a unique database of historical account level electricity and natural gas consumption (LA Energy Atlas) that is unprecedented in size within the U.S [10]. By integrating parcel level building attribute information with monthly account level energy consumption data, the LA Energy Atlas provides opportunities for the quantitative evaluation of how changes in the residential built environment relate to the sector’s energy consumption profile and its associated externalities [10,11]. Following these analyses, the study concludes with a Monte-Carlo sampling exercise that explore how energy savings attributable to the implementation of historical EE code changes compare to those which would have likely been achievable...
through constraints on historical rates of growth in new home size.

1.1. Background

California has the largest and most aggressive EE investment portfolio in the nation. In 2016 its total annual EE program budget was $1.4 billion. This is more than twice that of New York, the state with the next largest collection of EE programs [12]. According to the California Public Utility Commission (CPUC), the regulatory authority which administers California’s EE programs, 4% of this annual budget is dedicated to a combination of “field-based impact evaluations, market assessment, and other program evaluation activities” [13].

The CPUC’s EE program evaluation methods “estimate the potential energy savings for an energy efficient measure before it is installed based on predictions of typical operating conditions and baseline usage” [14]. An alternative to this approach are program evaluation methods which empirically observe historical changes in actual metered energy consumption. These types of empirical studies are fairly uncommon throughout the United States however, as consumer privacy regulations significantly constrain researcher access to account level historical energy consumption data [15,16].

1.2. Previous work

The majority of the previous related work comes from countries in Europe, Asia, and elsewhere where historical account level energy consumption data is more readily available [17–24]. While these studies are useful in the extent to which they inform the types of information which can be gleaned from the analysis of empirical consumption data, the direct transferability of their conclusions to the American context is somewhat limited due to the unique characteristics of our nation’s culture, economy, and building stock [25,26].

Of the similar U.S. based studies which do exist, many suffer from limited spatiotemporal data coverage and coarse levels of data aggregation [27–31]. For example, a comparable analysis conducted for New York City in 2012 used an energy consumption dataset made available by the Mayor’s Office of Long Term Planning and Sustainability [28]. The data disclosure agreement involved with this study stipulated that electricity and natural gas consumption records be aggregated to the zip code level. The net result of which was that only 191 aggregated consumption values were matched to the attributes of ∼1 million buildings within the city. In another recent U.S. based study, conducted by Kavousian et al. in 2013, account level smart meter data was obtained for California energy consumers within the Pacific Gas and Electric Utility’s (PG&E) service territory [29]. In this study, the scope of the analysis was similarly limited however, with the dataset comprising only 1628 accounts.

The most directly relevant previous work was an analysis conducted by Jacobsen et al. in 2013 [30]. This study used historical account level electricity consumption data to quantitatively assess the impact of a raft of new statewide building EE codes which had been introduced within the state of Florida in 2001. The work of Jacobsen et al. employed similar data integration procedures to those which were used to develop the LA Energy Atlas, including geocoding utility account addresses to the tax assessor parcel level. However, as the authors of this study note, after data validation filters had been applied, the final dataset that was used to conduct their analysis comprised a total of 64,471 residences. While significant in scale, this number only corresponds to about 5% of the sample size captured by the LA Energy Atlas. This issue of data coverage becomes a significant when considered relative to the fact that residential energy consumption values tend to be log-normally distributed. Consequently, large sample sizes are necessary to capture important observations which appear at the tails of the distribution.

2. Materials and methods

2.1. Data sources

The LA Energy Atlas includes monthly account level energy consumption data for all of the investor owned utility (IOU) service accounts and a portion of the municipally owned utility (MOU) service accounts located throughout LAC. The records used in this study therefore, corresponded to the subset of the single-family located homes within LAC for which both metered electricity and natural gas consumption data was available during the 2010 calendar year (n = 1,298,683) [10]. Through a multi-level process of parcel geocoding, the addresses associated with each of these individual utility accounts have been linked to building attribute information available from the LAC Assessor’s Office as well as Block Group level socio-economic data available from the U.S. Census Bureau’s 2010 American Community Survey [32]. The raw account level consumption values were originally provided in units of kWh/Month for electricity and therm/Month for natural gas. The values for natural gas have been converted from therms to kWh (1 therm = 29.3 kWh) to facilitate ease of comparison within the study’s various analyses.

The account level data contained within the LA Energy Atlas was obtained through agreements negotiated both with local MOU providers and, for the regional IOU provider, with the California Public Utilities Commission (CPUC). Table 1 lists the utility providers whose consumption data has been included in the LA Energy Atlas along with a breakdown of the relevant data access agreements involved.

Pursuant to the legally binding data privacy restrictions originally set forth by the CPUC in Rule-making Proceeding 08-12-009 and subsequently detailed in Decision 15-05-016, the raw account level electricity and natural gas consumption data underpinning this analysis can only be made public if suitably aggregated and anonymized [33]. These rules stipulate that for a set of residential utility ratepayer accounts, consumption data must be masked unless there are a minimum of 100 accounts contained within the set and no single account comprises greater than 15% of the set’s total aggregate consumption. All of the data and figures published as part of this analysis abide by these privacy requirements.

2.2. Descriptive analysis procedures

The first analysis quantitatively describes the depth and the breadth of the data sampled from the LA Energy Atlas with a figure that contains a series of bi-variate histograms constructed for several key co-variates (Fig. 1). All of the variables plotted, except for the home construction vintage, are depicted using a Log_{10} scale due to the wide spread of their distributions. Similarly, in all of the plots the count frequencies have also been Log_{10} scaled and are depicted using a gradient colormap (Blue-Green-Yellow). The broken lines (Red) and associated parameters describe the results of a simple linear model fit for each set of co-variates shown.

Table 1

<table>
<thead>
<tr>
<th>Utility</th>
<th>Type</th>
<th>Source</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern California Edison</td>
<td>IOU</td>
<td>Electricity</td>
<td>CPUC agreement</td>
</tr>
<tr>
<td>Southern California Gas Company</td>
<td>IOU</td>
<td>Natural gas</td>
<td>CPUC agreement</td>
</tr>
<tr>
<td>LA Department of Water &amp; Power</td>
<td>MOU</td>
<td>Electricity</td>
<td>Direct agreement</td>
</tr>
<tr>
<td>Burbank Water &amp; Power</td>
<td>MOU</td>
<td>Electricity</td>
<td>Direct agreement</td>
</tr>
<tr>
<td>Glendale Water and Power</td>
<td>MOU</td>
<td>Electricity</td>
<td>Direct agreement</td>
</tr>
<tr>
<td>Long Beach Gas and Oil</td>
<td>MOU</td>
<td>Natural gas</td>
<td>Direct agreement</td>
</tr>
</tbody>
</table>
2.3. Neighborhood analysis procedures

The second analysis exposes the spatial structure of median household income levels, home square footages, and combined energy intensities for single-family residential homes located throughout LAC (Fig. 2). Due to the aforementioned data privacy requirements constraining the geographic reporting of account level energy consumption data, the spatial unit selected for this analysis was the neighborhood level. This choice ensured that sufficient numbers of accounts would be aggregated together so as not to violate privacy constraints [33]. The neighborhood boundary dataset which was used comprises 272 distinct geographies and was generated as part of LA Times sponsored crowdsourced mapping effort [34].

The figure which was developed as part of this analysis contains a set of stem-plots which illustrate the degree to which median home sizes and home energy intensities within each neighborhood deviate from the LAC-wide median values. Within each stem-plot individual neighborhoods have been sorted along the horizontal axis by ascending magnitude of deviation and shaded according to a gradient colormap (Green-Yellow-Red) which corresponds to the percentile range of their magnitude of deviation. In all cases, standard deviations were used as the unit of measure.

2.4. Vintage analysis procedures

The third analysis quantifies differences in the median energy intensities of homes that were built during different construction vintage periods (Fig. 3). Visual inspection of the frequency counts of the construction vintages for the set of homes contained within the LA Energy Atlas revealed them to be non-uniformly distributed. This irregular distribution reflects historical patterns of residential property development within LAC [35]. In order to understand the significance of this periodicity in residential development on the structure of energy usage intensity we devised an algorithmic procedure to determine the temporal center and spread of the signal’s component peaks.

This procedure involves two key steps. The first uses an automated algorithm to identify discrete component peaks within the input dataset of construction vintage frequencies. The second uses the locations of the identified peaks to constrain the fit a Gaussian Mixture Model (GMM). This GMM fitting procedure approximates the shape of the vintage frequency signal through the additive combination of a set of discrete

---

Fig. 1. Bivariate histograms of LAC single-family home (A) square footages relative to construction vintage years, (B) electricity usage intensities relative to construction vintage years, and (C) natural gas usage intensities relative to construction vintage years.

Fig. 2. (A) Stem-plot showing the extent to which the median home square footages within each neighborhood deviate from the LAC-wide median. Stems have been sorted along the horizontal axis by the degree to which each neighborhood’s median household income deviates from the LAC-wide median. (B) Stem-plot showing the extent to which the median combined home energy intensities within each neighborhood deviate from the LAC-wide median. Here again, stems have been sorted along the horizontal axis on the basis of neighborhood income deviations. Four neighborhoods with notable combinations of EUI and square footage deviations have been highlighted.

---

The third analysis quantifies differences in the median energy intensities of homes that were built during different construction vintage periods (Fig. 3). Visual inspection of the frequency counts of the construction vintages for the set of homes contained within the LA Energy Atlas revealed them to be non-uniformly distributed. This irregular distribution reflects historical patterns of residential property development within LAC [35]. In order to understand the significance of this periodicity in residential development on the structure of energy usage intensity we devised an algorithmic procedure to determine the temporal center and spread of the signal’s component peaks.

This procedure involves two key steps. The first uses an automated algorithm to identify discrete component peaks within the input dataset of construction vintage frequencies. The second uses the locations of the identified peaks to constrain the fit a Gaussian Mixture Model (GMM). This GMM fitting procedure approximates the shape of the vintage frequency signal through the additive combination of a set of discrete
Gaussian components; each representing a discrete construction boom period.

Two key parameters were used to constrain the automated peak fitting algorithm. The first was an amplitude threshold which sets the minimum difference in amplitude that separates any two distinct peaks. The second was a distance threshold which sets the minimum distance between any two distinct peaks. In order to select the combination of these two parameters that delivered a final GMM fit which best approximated the input signal with the minimum number of component peaks, a grid-search was conducted which iteratively generated individual GMM fits for each set of peaks derived from 180 different pairwise combinations of the two input hyper-parameters. From this collection of candidate models a single, optimal GMM fit was selected as the model with the minimum Akaike information criterion (AIC) score. AIC scores are useful for this type of model selection exercise as they combine measures of goodness of fit with measures of model complexity to favor a model with the best performance but also the simplest structure.

This optimal GMM fit decomposed the input construction vintage frequency time series into a set of nine non-uniformly spaced vintage periods (Fig. 3). The boundaries between each period are defined by the intersection of each successive pair of GMM component peaks. Using these nine component construction vintage frequency periods, a set of pairwise two way t-tests were performed to statistically compare the potential significance of the differences of the EUI levels of homes built within each vintage period.

2.5. Policy analysis procedures

The fourth and final analysis relates the findings of the previous analyses to the conservation efficacy of long standing California residential building EE codes. In 1978 the California State Legislature adopted the California Energy Code, Part 6 of the California Building Standards Codes, which is Title-24 of the California Code of Regulations [36]. These regulations mandated escalating annual improvements in the code-minimum EE standards for various building shell components and energy appliances. The date of Title-24’s introduction provides as a useful benchmark for a longitudinal analysis of California’s energy policy performance.

Title-24 measures have no doubt played a significant role in stimulating energy conservation among LAC’s single-family residential housing stock in the years since they were first introduced. However, given the significant growth in median home sizes which has occurred since they were introduced, we sought to evaluate whether or not it
would have been possible to achieve the same energy conservation gains using an alternative set of policy measures that were instead focused on controlling the growth in the size of new homes over time. In order to answer this question we developed a Monte-Carlo sampling procedure to evaluate the likely conservation outcomes associated with the following three scenarios:

Scenario 1. If Title-24 standards had not been implemented (pre-1978 EUIs) and home size growth continued, as it has, unabated.

Scenario 2. If Title-24 standards had not been implemented (pre-1978 EUIs) but home size growth had been constrained to pre-1978 levels.

Scenario 3. If Title-24 standards had been implemented in conjunction with the constraint of home size growth to pre-1978 levels.

As part of this Monte-Carlo sampling procedure we divided the full set of homes sampled from the LA Energy Atlas database into two groups: the first group consisted of 1,076,008 homes which were built between 1900 and 1978, the year that Title-24 measures took effect. The second group consisted of 222,675 homes which were built between 1979 and 2010. Within each group, the combined EUI and square footage values of homes were first \( \log_{10} \) transformed then separately fitted to a reference Gaussian distribution model. The parameters of these reference Gaussian distribution model fits were then used to inform the Monte-Carlo sampling procedure. Each sample was comprised of a set of synthetic realizations of homes built in the post-1978 vintage period. A single sample run involved the repeated random selection of values from one of three contrived bivariate distributions. These bivariate distributions were assembled such that their mean vectors (\( \mu \)) and covariance matrices (\( \Sigma \)) reflected the different pairwise combinations of the pre and post Title-24 home groups’ attributes associated with the three policy scenarios.

According to this procedure a single realization therefore consisted of a correlated set of values for the total square footage and combined EUI (electricity and natural gas) for a single synthetic home. Multiplying these values together gave that home’s expected total annual energy consumption. Within each Monte-Carlo sample, individual homes’ annual totals were then summed to give an expected value for the total combined annual energy consumption across all homes within the post Title-24 group. A total of 1 million sampling runs were evaluated for each of the proposed scenarios.

3. Results

3.1. Descriptive analysis results

The average annual growth rate in the median size of single-family homes within LAC was roughly 0.86%/Vintage Year for the 1900–2010 vintage period (Fig. 1A). By comparison, the annual rates of decline in electricity and natural gas consumption intensities per square foot were observed to be \(-0.23\% / \text{Vintage Year}\) and \(-0.69\% / \text{Vintage Year}\), respectively (Fig. 1B &C).

3.2. Neighborhood analysis results

The overall trend among LAC neighborhoods is one of negative...
correlation between home size and combined EUI (Fig. 2A & B). Several neighborhoods were observed defy this overall trend, in terms of possessing median home square footages and median EUI’s which were both above countywide averages. From an energy conservation standpoint, this combination of characteristic is undesirable. The identity of these neighborhoods and their corresponding EUI/square footage deviations have been highlighted within the figure.

3.3. Vintage analysis results

Significant reductions in median natural gas intensity by construction vintage year were observed during the transition from ($P_{pre}$) to ($P_{post}$) vintages (Fig. 3A–C). Progressive declines continued over successive vintage periods until ($P_f$) after which there was a single period of increases ($P_i$), until the final vintage period ($P_f$) when median natural gas intensity reductions were resumed. Differences in electricity intensity between vintage periods were noticeably less pronounced across all vintage years (Fig. 3A & D-E). This decoupling of electricity intensity from consumption vintage is likely attributable to the rapid replacement cycle of electrical appliances and plug loads when compared to the more durable performance characteristics of a building’s thermal shell [37].

3.4. Policy analysis results

Two-way Kolmogorov-Smirnov tests reveal that the distribution of home energy intensities and home square footages are significantly different between the pre-Title-24 and post-Title-24 groups at the 95% confidence level (Fig. 4A & B). For the first scenario, Monte-Carlo sampling results indicate that in the absence of Title-24 EE regulations, for the year 2010, total combined energy consumption for the sample of post-1978 vintage homes would likely have been $\sim 1.9 \text{TWh}$ ($\pm 22.9\% \pm 0.5\%$) higher than observed levels (Fig. 4C). This result confirms the significance of the conservation impacts associated with the implementation of Title-24 EE measures. For the second scenario, in which Title-24 EE gains were effectively eliminated but the distribution of home square footages were constrained to pre-1978 levels, Monte-Carlo sampling results suggest that observed total combined energy consumption could have been further reduced by $\sim 1.1 \text{TWh}$ ($\sim 17.1\% \pm 0.7\%$) below observed levels (Fig. 4C). Finally, for the third scenario, Monte-Carlo sampling results illustrate the beneficial effect of combining EE measures designed to reduce energy intensities per square foot with parallel constraints on growth of the size of new homes. Under this scenario, total combined energy consumption was expected to be reduced by $\sim 2.4 \text{TWh}$ ($\sim 37.5\% \pm 0.9\%$) below observed levels (Fig. 4C).

4. Discussion

In 2015, California state legislators passed Senate Bill 350 (SB 350) which mandates a 40% reduction in statewide GHG emissions from 1990 levels by the year 2030 [38]. As part of the envisioned pathway for achieving this ambitious goal, SB 350 requires a doubling of the EE of electricity and natural gas end uses statewide over the same time period. This study’s findings for LAC cast significant doubt over the feasibility of achieving SB 350’s GHG emissions reductions targets through increases in end use EE alone. Rather, the data suggests that our collective approach to the problem of increasing energy conservation must adopt a broadened perspective; one which accounts for parallel and confounding trends in the development of the state’s demographics and residential built environment [39]. These are findings which have been echoed in other previous studies which have separately sought to quantify the GHG mitigation potentials for the U.S. residential sector at large [40].

We believe that a new generation of energy policies should be developed which shift the emphasis from efficiency to conservation. Such policies should necessarily be implemented differently for households at opposing ends of the income spectrum and tied to rigorous empirical evaluation methods. For example, programs which provide financial assistance to help overcome the high up-front costs of deep home efficiency retrofits and major appliance upgrades should continue to be made available to lower income households. However, among higher income households, which do not need the same financial assistance, new policy measures should be devised which limit the continued growth in the average size of newly constructed homes. Precedents for this type of policy intervention, sourced from both foreign and domestic contexts, most commonly take the form of floor to area ratio (FAR) restrictions which have been implemented relative to specific residential property classifications [41-43].

5. Conclusions

Throughout LAC, homes which were located in lower income neighborhoods tended to be both older and be smaller in size than homes located in higher income areas. While such a finding may have been anticipated, this study reveals important new information regarding the relative magnitude of these differences. Specifically, recent growth in the size of newer homes, which are predominately located in higher income neighborhoods, has occurred at a rate that is 60% greater than the corresponding rate of combined EUI reduction. The disproportionality of these rates indicates that a large fraction of the energy savings that would have been expected from recent residential EE improvements were likely lost as a result of parallel growth in the sheer size of newer homes.

The large range of median natural gas intensities which were observed among homes built in different vintage periods suggests that significant energy efficiency opportunities still exist among older vintage homes located within lower income communities. As a result, the implementation of EE measures designed to improve building shell thermal performance would likely result in substantial energy savings within major urban areas such as LAC [44-46]. The significant economic cost associated with implementing these types of deep EE retrofits has thus far been a major hurdle to their more widespread adoption, however. These high retrofit costs serve to reiterate the profound importance of investing in residential building thermal shell performance at the time of construction as a means of enhancing long term energy conservation potential.

EE programs remain an important part of ongoing efforts to reduce residential building sector energy use. Many of these programs can also provide additional benefits such as lowering consumer energy bills and enhancing levels of thermal comfort within homes. However, as the results of this analysis have demonstrated, even with the implementation of thoughtful and robust building EE standards, such as those associated with California’s far reaching Title-24, there remains a need to address the persistent growth in the size of new residential buildings. This growth, if continued unabated, will significantly constrain the feasibility of achieving the State’s broader GHG emissions reductions goals.

Funding sources

This work was supported by funding from the County of Los Angeles and the California Energy Commission (CEC) [GFO-15-312].

Acknowledgements

The authors would like to acknowledge the work of analysts Hannah Gustafson and Dan Cheng in the initial development and continued expansion of the LA Energy Atlas database.