Energy Policy 64 (2014) 263–272

Contents lists available at ScienceDirect

Energy Policy

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

How much information disclosure of building energy performance is necessary?

David Hsu*

University of Pennsylvania, 210 S. 34th Street, Philadelphia, PA 19104, United States

HIGHLIGHTS

- A comprehensive panel dataset of energy performance and building characteristics was assembled and cleaned.
- The effectiveness of the disclosed information to predict building energy performance was tested using a regression model.
- Building-level variation has a greater effect than any building characteristic or systems.
- Benchmarking data alone predicts energy performance equally as well as both benchmarking and engineering audit data together, and better than audit data alone.

ARTICLE INFO

Article history:
Received 15 May 2013
Received in revised form
14 August 2013
Accepted 28 August 2013
Available online 8 October 2013

Keywords:
Building energy performance
Disclosure laws
Analysis of variance
Benchmarking
Auditing

ABSTRACT

Many different governments have begun to require disclosure of building energy performance, in order to allow owners and prospective buyers to incorporate this information into their investment decisions. These policies, known as disclosure or information policies, require owners to benchmark their buildings and sometimes conduct engineering audits. However, given substantial variation in the cost to disclose different types of information, it is natural to ask: how much and what kind of information about building energy performance should be disclosed, and for what purposes? To answer this question, this paper assembles and cleans a comprehensive panel dataset of New York City multifamily buildings, and analyzes its predictive power using a Bayesian multilevel regression model. Analysis of variance (ANOVA) reveals that building-level variation is the most important factor in explaining building energy use, and that there are few, if any, relationships of building systems to observed energy use. This indicates that disclosure laws requiring benchmarking data may be relatively more useful than engineering audits in explaining the observed energy performance of existing buildings. These results should inform the further development of information disclosure laws.

1. Introduction

Buildings have been recognized as a key pathway and setting for the consumption of energy and carbon emissions worldwide (Mazria, 2003; Pacala and Socolow, 2004; Ürge-Vorsatz and Novikova, 2008). In the United States, 40% of all energy is used by residential and commercial buildings alone (USEIA, 2012). Estimates of the portion of total energy used by buildings range from 20% to 40% worldwide, and is expected to grow rapidly in future years (Pérez-Lombard et al., 2008).

In building energy efficiency, as in other sectors, over the past forty years there has been ongoing debate about the existence or nature of the "energy efficiency" gap. Information plays a key role in this debate, such as whether people have information about energy efficiency investments; whether they pay attention to it; or if it is sufficient and salient enough on which to act. Research on this gap has therefore focused on the reasons why owners and occupiers may be unable or unwilling to invest in energy efficiency, even when it is considered to be rational based on the returns predicted by engineering-economic models (see, for example, Allcot and Greenstone, 2012; Blumstein et al., 1980; Jaffe and Stavins, 1994; Levine et al., 1995). Lack of information contributes to other structural barriers to energy efficiency such as information asymmetry, bounded rationality, and uncertain risks and rewards.

The emphasis on the role of information is clearly reflected in recent efforts to mandate energy performance disclosure, which are intended to transform the market for energy efficiency in buildings. Mandatory disclosure policies have built on previous efforts to establish voluntary benchmarking schemes in many countries (Burr et al., 2011). These policies have a number of

* Tel.: +1 215 746 8543.
E-mail address: hsuyd@design.upenn.edu

0301-4215
http://dx.doi.org/10.1016/j.enpol.2013.08.094

attractive features. They have relatively low implementation costs compared to other energy efficiency policies. And, in contrast to other policies that require legislation at the federal level in the United States, they can be passed at the local and state level. At present, the states of California and Washington, and eight major cities, including Washington DC, Austin, New York, Seattle, San Francisco, Philadelphia, Minneapolis, and Boston have passed such disclosure laws. Chicago is currently deliberating such a policy, and other numerous other cities have expressed interest. In the European Union, the United Kingdom has implemented Energy Performance Certificates for domestic buildings, Display Energy Certificates for public buildings.

These disclosure efforts also play an important role in creating new information for owners, policymakers, and analysis. In just the past few years, cities that have passed these policies plan to gather information on more than 60,000 buildings, more than 4 billion square feet of commercial and multifamily space, or more than 4% of all U.S. commercial real estate (Florance et al., 2010). This is considerably larger than existing surveys of commercial or residential buildings which are conducted by the U.S. Energy Information Administration at five year intervals, and sometimes less frequently. This information is also particularly important because it addresses actual building performance, which many studies indicate can differ from the performance predicted from the design, modeling, or certification processes (Norford et al., 1994; Turner and Frankel, 2008; Pang et al., 2012). Furthermore, occupant behavior and building commissioning have both been found to play critical roles in determining energy use, leading to a renewed emphasis on empirical validation of building models and operational efficiency measures (Judkoff and Neymark, 2006; Mills, 2011).

Disclosing information about energy performance, however, requires a non-trivial amount of work necessary by all parties involved, including policymakers, building owners, tenants, and investors. Information needs to be gathered, verified, structured, analyzed and disclosed before anyone can act on it. Policymakers, building owners, consultants, and academics are only at the beginning stages of developing new policies for information disclosure, and need to strike an appropriate balance between cost and effectiveness. Given wide variation in the laws that have been passed by various jurisdictions and their initial experiences, a critical question therefore arises: how much and what kind of information should be disclosed about building energy performance, and to what purposes?

This paper answers this question by first assembling a comprehensive dataset of energy consumption and building characteristics for a group of multifamily buildings in New York City, similar to the best information that could be gathered from a range of existing policies. It will be discussed further below why New York City multifamily buildings serve as a test case for analysis. This paper then evaluates the effectiveness of three different information disclosure policies, as implemented in New York, Seattle, and San Francisco respectively, in predicting building energy performance. Each city's policy requires a different level of information disclosure. By modeling the same dataset through the lens of each policy, and using regression models to explain the observed energy performance, this paper tests whether the required levels of information disclosure for each city could enable market transformation by allowing other actors in the marketplace to assess the energy performance of buildings. Analyzing one comprehensive dataset using policies that are similar in intent but different in implementation enables this paper to suggest improvements in policy design.

The remainder of this paper is organized as follows. Section 2 reviews the literature about information and disclosure policies. Section 3 describes the policy context, and discusses variations between the policies of New York, Seattle, and San Francisco. Section 4 describes the methods used to analyze the data and how this analysis would be viewed differently from the perspective of each city's policy. Section 5 describes the dataset, and Section 6 interprets the results. Finally, Section 7 concludes the paper by discussing the findings and limitations of the research.

2. Literature review

This section describes connections between energy efficiency, building performance, and information disclosure policies in the academic literature.

Markets require information to work, and lack or asymmetry of information has long been recognized as a potential barrier to investments in energy efficiency in buildings. Owners are unlikely to invest in energy efficiency if there is uncertainty about the future benefits of energy efficiency relative to upfront investments. Furthermore, if prospective buyers cannot evaluate and compare the energy performance of buildings, then there is no way for owners to capitalize the value of future energy savings. Everyone seems to agree that more, better and shared information would decrease investment inefficiencies, reduce adverse selection, and enable better targeting of energy efficiency policies to increase their overall effectiveness in terms of cost and energy reductions (see, among many others, Akerlof, 1970; McKinsey, 2009).

Information programs to assist environmental markets can be generally classified into two broad categories: product labeling and reporting requirements (Stavins, 2003). Product labeling has been extensively used for appliances through programs such as Energy Star (Banerjee and Solomon, 2003), and has directly led to the numerous building rating and labeling programs in use today, such as Leadership in Energy and Environmental Design (LEED); the European Union's Energy Performance of Buildings Directive, and Australia's building certification scheme. Several academic studies find significant premiums associated with Energy Star and LEED certifications (Eichholtz et al., 2009; Fuerst and McAllister, 2011; Miller et al., 2008). In contrast, reporting requirements are relatively new for building energy performance, compared to the broader environmental economics literature. This is in part because it has been, until recently, relatively difficult to obtain data about energy use in large samples of buildings.

Disclosure of building operating performance fits into a general trend of policies that have been applied in other areas, with varying degrees of success. Disclosure laws go by many names, including but not limited to, information laws, regulatory disclosure, or transparency laws. Weil et al. (2006) defines this growing area as “the mandatory disclosure of information by private or public institutions with a regulatory intent.” Such laws have been applied in many policy areas, including financial markets, health care and nutrition, workplace hazards, sex offenders, and corruption, with mixed results depending on their design and context (Fung, 2007; Rosenthal, 2012; Winston, 2008).

One example, the Toxic Release Inventory data by the U.S. Environmental Protection Agency, illustrates some of the issues raised by the quantity and quality of the data in disclosure laws. While some studies find that the release of the information by the US Environmental Protection Agency leads to positive outcomes such as better environmental behavior by firms (Konar and Cohen, 1997) and increased allocations to environmental and natural resource programs (Patten, 1998), other studies also find that there can be unintended consequences, such as inefficient allocation of benefits and costs for relatively clean and dirty firms (Delmas et al., 2010). Fewer studies are able to get at the impact of the quality of the information itself. Two notable exceptions include Bae et al. (2010), who find that states that engage in efforts to analyze and process the data achieve lower toxic risks than those that simply release the data; and Massey (2011), who uses surveys and interviews to identify the most effective aspects of Massachusetts' Toxic Use Reduction Act (TURA) program.

Building energy performance disclosure laws can also be clear and effective because they are focused on a single metric – observed energy consumption – which can be obtained from a relatively small number of fuel providers and utility companies. In Weil et al. (2006), the concept of 'embeddedness' for information users and disclosers is used to conceptualize how these disclosure laws may overcome barriers from the energy efficiency literature. In order for disclosed information to be effective, both users and disclosers must make use of accurate and valid information in a repeated cycle of communication and action. Users are connected to their buildings in a number of ways: financially, as owners, tenants, or prospective buyers; spatially, as utility planners, tenants or local community groups; or through the jurisdiction of government and markets, in which users may benefit from high-quality, aggregated information about building energy use in the local industry or market. Disclosers of energy data, usually building owners or utilities, need to have a direct economic incentive to better position efficient buildings relative to other less efficient buildings. Both users and disclosers may discover new ways to save energy when comparing themselves to peers in the aggregated data.

Better understanding of how information disclosure is used should also impact how policies are formulated and implemented in a number of ways. First, Allcott and Greenstone (2012) note that there may be significant heterogeneity in the return on investments in energy efficiency. Identifying good investments is a necessary first step to better design and target energy efficiency policies. Second, building energy disclosure laws will greatly assist movement towards performance-based energy codes, rather than those based on prescribed efficiency measures (Hewitt et al., 2010). Third and finally, disclosure to municipal governments may lead to improved energy efficiency through better understanding of the building stock, and subsequent improvement in policy design by governments.

How much or what kind of information is needed has not been addressed very often in the building energy efficiency literature. Although some authors have suggested both labeling and energy efficiency audits as a policy to overcome the barrier of asymmetric information (Hirst and Brown, 1990; Levine et al., 1995), empirical studies of the effectiveness of energy efficiency audits are relatively few in number. In the industrial sector, Anderson and Newell (2004) find that government-sponsored energy audits do encourage industrial users to implement energy efficiency projects, while Schleich (2004) finds that only some auditors actually result in lower energy consumption. Surprisingly, considering the emphasis placed upon audits in energy efficiency policy, this author could find no academic or peer-reviewed studies of the effectiveness of energy audits on building energy consumption. The closest that the author could find was two recent studies which consider how residential homeowners and auditors exchange and act on the kinds of information typically included in audits (Palmer et al., 2012; Ingle et al., 2012). Shapiro (2011) is a general review of the problems that can lead to bad energy audits, including inadequate review and analysis, overestimated saving and/or low or missing cost estimates, and no costs based on life-cycles or improvement life. Krarti (2011) lists several verification methods for energy audits, including regression and time variant models.

Finally, it is again important to emphasize that audits as energy information may have a very different effect on owner-occupiers, as opposed to a tool for communicating in the market. While many measurement and verification studies have been done in the California utility industry, such as evaluations of audits in residential buildings (Robert Mowris and Associates, 2008), commercial buildings (Robert Mowris and Associates, 2007), affordable housing (KEMA, Inc., 2006), and school buildings (Itron, Inc., 2006), most of these studies are intended to measure the impact of audits on energy use by owners or occupiers. None of these studies address how information disclosure may affect the valuation of energy efficiency in buildings by prospective owners or tenants.

3. Policy context

Small changes in implementation can make a large difference in the effects of a policy. Evaluating what kind, and how much information, assists in the comparison of buildings therefore has important implications for the further development of energy disclosure policies. In order to achieve the goal of maximizing social welfare, policymakers should balance the benefits and costs of gathering this information within a theory of policy action, and analyze the effectiveness of these policies in achieving these goals. Interpreting this in terms of energy performance disclosure laws, the minimum criterion would be that policymakers should seek to require information disclosure at the minimum cost that is required to enable an increase in benefit or welfare for each individual (Zerbe and Dively, 1994). Another criteria may be that there should be a net social benefit to information disclosure, although benefits and costs will not be allocated evenly.

Of the major cities, at the time of writing, only four have collected data: Austin, New York, Seattle, and San Francisco. All of these ordinances have different requirements for how the information will be gathered and disclosed. While all of the ordinances require utility data to be disclosed to the city government, some cities (New York and San Francisco) make data available on a public website, while other cities (Austin and Seattle) make data available only to prospective buyers at the time of transaction. Legislation for these policies also often requires the city to produce a report on trends in energy use in the major building sectors and to analyze the quality of the data (see, for example, Kerr et al., 2012, to which this author contributed). The data requirements for the cities vary greatly, and can generally be divided into two categories: benchmarking and audit requirements. The City of Seattle only requires benchmarking disclosure, while Austin, New York, and San Francisco all require additional energy auditing to varying degrees.

Benchmarking laws typically require building owners simply to gather their utility data, usually from monthly bills, and to report some basic information, such as building characteristics and size. The U.S. Environmental Protection Agency's Portfolio Manager website is often used to collect this information, and often the data is benchmarked in terms of total energy use or energy use intensity (EUI), which is total energy use normalized by total gross floor area. The Portfolio Manager website also provides an Energy Star rating for commercial buildings (not yet for multifamily), as well as weather-normalization of total energy use and EUI.

Audit laws, such as in Austin, New York and San Francisco, require a higher level of information gathering and disclosure. In the United States, a comprehensive energy efficiency audit requires a engineer licensed by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) to perform measurements and calculations at several different levels. An ASHRAE Level 1 audit is a walk-through survey, while an ASHRAE Level 2 audit requires a more comprehensive energy survey and data analysis. The highest level, an ASHRAE Level 3 audit, requires detailed analysis of possible capital intensive-modifications, including modeling and simulations (ASHRAE, 2011). Again, each city has different information requirements. Table 1 summarizes each city's laws, including data passed, total square footage, and

Table 1
Table of benchmarking and audit legislation.

City	Legislation/ordinance	Benchmarking requirement, frequency	Benchmarking disclosure	Auditing requirement, frequency	Auditing disclosure
New York	Local laws 84, 87	Annual disclosure of total energy use and building characteristics	Public	ASHRAE Level 2 for all buildings every ten years	To City only
San Francisco	Energy Performance Ordinance	Annual disclosure of total energy use and building characteristics	Public	ASHRAE Level 1 or 2 for each building every five years	Public
Seattle	CB 116731	Annual disclosure of total energy use and building characteristics	Only to potential buyers and tenants	None	None

information requirements. The City of Austin requires both audits and disclosure, while the City of New York will require their first ASHRAE Level 2 audits by the end of 2013, with the entire population of buildings audited every ten years on a rolling basis. The City of San Francisco requires a ASHRAE Level 1 or 2 audit every five years in order to certify the energy benchmarking, and to provide a list of possible energy efficiency improvements.

In this paper, the effect of auditing and benchmarking is considered using only multifamily buildings in the City of New York for the following reasons. The City of New York is the largest of the cities engaged in benchmarking and auditing disclosure, comprising nearly 50% of all benchmarked buildings in the United States alone. The City of New York also had extremely good compliance and take-up of its benchmarking policies (over 75%), compared to the first year in Seattle and San Francisco. Use of a single comprehensive dataset, which is discussed further in Section 5, allows comparison of different laws applied to the same data without introducing between-city variation, as would be the case with different datasets from different cities. Finally, because the cities are all applying benchmarking with Portfolio Manager and auditing with ASHRAE Level-2 audits, insights about policy implementation from the City of New York should be generalizable to the other cities, at least in the United States.

The key difference between benchmarking and audit policies is that the former provides a snapshot of building-level variation in energy consumption only, while the latter is expected to provide critical engineering information about building characteristics, systems, occupancy, and space uses which will help to explain these differences. It is possible that these two types of information may have very different effects in the marketplace. Benchmarking reduces building energy use to a single metric (total weather-normalized EUI) while auditing provides a much more detailed portrait of building characteristics and may include suggested energy efficiency improvements.

Finally, and very importantly, cost differs significantly between benchmarking and auditing. Consultants charge more for more complex modeling and data analysis, often by a factor of 10 (California Energy Commission, 2000; U.S. Department of Energy, 2011). To take a recent example, benchmarking in New York City was estimated to cost approximately $500-$1500 USD per building, while auditing is significantly more expensive at approximately $1.50 per square meter (Kerr, 2013). For a typical New York City building of approximately 20,000 square meters, the difference between benchmarking and auditing in the same year computes to almost $30,000 USD, though under City of New York's law, audits are only required once every ten years.

4. Methods

This paper applies a general regression model in order to evaluate the systematic relationships that could be inferred from building level audit and energy consumption data, in order to estimate building-specific effects, and to analyze variance components. By including all of these effects in the model, and then including or removing various parts of the data in the model, we can predict what comparisons could be made with different combinations of benchmarking and audit data. Model 1 is fit using the full range of consumption and engineering information and is similar to what could be learned from New York's or San Francisco's ordinances. Model 2 is fit only using weather-normalized data is similar to what could be learned from Seattle's benchmarking-only ordinance. Model 3 is fit using only information about the building systems, similar to what could be learned in an engineering audit, and is called an "engineering-only" model.

The model structure applied in all three cases is a multilevel, two-way, crossed model with mixed effects. A two-way crossed model is used because there are clearly building-specific and seasonally specific effects in the total energy consumption of each building in a given month; estimation of this model is similar to estimating panel-corrected fixed effects and standard errors (Wooldridge, 2002). Bayesian methods were used to estimate the model for the following reasons:

- Multilevel modeling allows flexible specifications of hierarchical data, such as in studying the effects of individual building characteristics as well as to analyze components of variance (ANOVA) (Faraway, 2005; Gelman and Hill, 2007).
- Mixed effects are used to specify randomly varying intercepts for each building and month separately, as suggested by Gelman and Hill (2007). The main variable of interest is the variance in the random effects themselves, because they describe the variation in the average energy consumption between buildings.
- Classical methods for panel datasets and ANOVA do not estimate unbalanced datasets well, i.e., datasets where there are varying number of observations for each building or where there is missing data, because inference depends on assumptions about the degrees of freedom and the number of observations, while Bayesian methods are relatively robust to unbalanced datasets (Gelman, 2005).

The dependent variable in the analysis is energy consumption normalized by area, as measured by energy use intensity (EUI). The EUI is to be explained by the independent variables representing building characteristics, because buildings with different physical characteristics can be expected to have different levels of energy consumption (U.S. Environmental Protection Agency, 2011; Chung, 2011; Eddy and Marton, 2012).

The general model specification was as follows. As indicated by the square brackets, each observation $i = 1 \ldots N$ belongs to the building $j = 1 \ldots J$ and in the calendar month $t = 1 \ldots 12$. The equation form of the statistical model is parameterized with the following structure:

$$y_{i[jt]} = \mu + \beta_k x_{ki[jt]} + \theta_j + \theta_t + \epsilon_{i[jt]} \quad (1)$$

$$\theta_j \sim \mathrm{Normal}(0, \sigma_j^2) \quad (2)$$

$$\theta_t \sim \mathrm{Normal}(0, \sigma_t^2) \quad (3)$$

$$\epsilon_{i[jt]} \sim \mathrm{Normal}(0, \sigma_y^2) \quad (4)$$

where $y_{i[jt]}$ is the observation i of the log-transformed energy use intensity (EUI) for a given building j in month t. EUI in terms of kilowatt-hours per square meter (kWH per m^2) is used as the observed outcome to be explained, in order to normalize total energy use by building size, and the log-transform is used because of the extreme left skewness of the EUI distribution. Individual predictors are described by $i \times k$ dimensional matrix $x_{ki[jt]}$ for a given building j or month t, such as a given building characteristic or energy system for a particular building, or the average temperature or total precipitation for a given month, respectively.

The model parameters therefore to be estimated are μ, which is the overall model intercept, or grand mean; β_k, which are the coefficients of the individual predictors; and the building- and time-specific group effects, θ_j and θ_t, respectively. Group effects are specified in Eqs. (2) and (3) as normally distributed with separate group-level standard deviations σ_j and σ_t, and zero means to avoid identification problems with the grand mean. Individual errors $\epsilon_{i[jt]}$ are specified in Eq. (4) as a normal distribution with zero mean and a super-population standard deviation of σ_y.

Model estimation was carried out using Gibbs sampling, as a special case of Markov chain Monte Carlo (MCMC) simulation. Gelman and Hill (2007) provides an excellent general introduction to these methods for various data structures, based on work by Dempster et al. (1981), Efron and Morris (1975), Gelfand and Smith (1990), Lindley and Smith (1972), and Pauler et al. (1999). The R language was used to conduct the analysis including JAGS software and the Rjags package (Plummer, 2012, 2013; R Development Core Team, 2012).

Model fit is assessed using proportion of explained variance, or R^2, or the ratio of the variance of the residuals ($\epsilon_{i[jt]}$, or $y-\overline{y}$) and variance of the predictors (θ). The data-level predictors include the linear terms $\beta_k x$, as well as the building and monthly specific random effects, θ_j and θ_t.

$$R^2 = 1 - \frac{\mathrm{E}\left(\sum_{i=1}^{n}(y-\overline{y})^2\right)}{\mathrm{E}\left(\sum_{i=1}^{n}(\theta-\overline{\theta})^2\right)}$$

This metric has the conventional interpretation: R^2 is a percentage, and as it nears 1, more of the variance in the data is explained by the model and predictors.

5. Data and results

This section describes the assembly and cleaning process for the dataset, and the results of the modeling approach. Billing data and building typology data was provided by FS Energy, an energy advisory subsidiary of First Service Residential, a major real estate management firm that manages a large number of multifamily buildings across North America and in New York City.

Utility billing data was assembled as a result of a comprehensive data integrity effort carried out by FS Energy. This process first identified gaps in utility bill records for buildings, and obtained all missing records for buildings from utility providers and energy service companies (ESCOs) if they were involved. Overlapping bills were resolved and duplicate bills were removed. As a result, the utility billing data for each building is a comprehensive record of monthly energy consumption and does not include any imputed or estimated consumption figures (Mehta, 2013).

The existing utility data, which has starting and end points on random calendar days, was then recast into monthly bills which coincide with calendar months. This was achieved by first converting the monthly consumption figures into a daily average for each day of the period between the start and end date for each bill, and then re-aggregating on all bills to get consumption for each fuel type by building in each calendar month. Although the utility data contained bills for different fuel types, the vast majority of information was for either electric or natural gas, with relatively fewer bills for steam and heating oil of all types. This monthly energy and water data was then joined to climatic information obtained from the U.S. National Oceanographic and Atmospheric Administration Climatic Data Center, including daily maximum and minimum temperatures, heating-degree-days, and cooling-degree-days.

The building typology data file contains information on 361 buildings in New York City. This data was all gathered by a team of company-employed and certified auditors using a consistent auditing methodology, which is unusual for a large group of buildings. The audits were carried out at a level equivalent to a Level-2 ASHRAE audit. The building typology data includes key typology information that would be expected to have an impact on total building energy consumption, such as key information about building systems for heating, ventilation, and air-conditioning, as well as information about space uses, including commercial spaces and additional special building features such as pools, garages, and elevators. Previous studies compared these buildings to the general population of New York City multifamily buildings and found them similar in distribution of EUI and size (Kerr et al., 2012).

It is also important to discuss the limitations of the building typology data. The building typology data contains both direct-metered and master-sub-metered buildings, which could have significantly different building profiles, because some of the typology and billing data represents full buildings in some cases and only common areas in others.

The data also required cleaning in several different ways. Categorical responses that were inconsistently coded but in which the meaning was clear (i.e. “zero” versus “0” versus “none”) were recoded. If a particular numerical data point was ambiguous and unclear, then a zero was entered to omit this data point from the overall regression model. Many of the key engineering variables were coded simply as binary variables, such as whether or not a building control system was indicated, or if there are heating problems. Binary variables were used, because if they do not show significant results, it is unlikely that more detailed categorical data will show significant results either. In cases where some surveys consistently showed more detailed information, such as specific ages of particular equipment, continuous variables were interacted with binary variables. Audits were only conducted once in each building, so major building changes could not be identified in order to explain changes in energy consumption time series, though this happened rarely.

The panel dataset was created by joining the utility bills and building typology files on unique building identification codes. Once this data was joined, there were only 255 buildings that had consistent billing records in the three year period from 2008 to 2011 for all fuel types including electricity, natural gas, heating oil, and steam bills. All energy figures were then converted to kilowatt-hours using regional conversion factors and added together to get total building energy consumption, and then divided by total building areas to get EUI.

It is also worth comparing this dataset to other recent studies, such as Belzer (2009), Howard et al. (2012), and Ryan and Sanquist (2012). This dataset contains a well-verified monthly time-series of energy consumption on existing individual buildings over a relatively long period over three years, and matches it to a building typology file.

Table 2
Descriptive statistics for monthly bill data, 2008–2011.

	Variable	Mean	Std dev	Min	q25%	Median	q75%	Max
1	Average temperature (°C)	13.4	8.7	−2.3	5.8	12.9	21.7	27.4
2	Total precipitation (mm)	122.4	76.8	23.6	76.0	105.2	149.0	481.3
3	Heating degree-days (to 18 °C)	60.7	84.8	0.0	0.0	8.1	112.3	291.6
4	Cooling degree-days (to 18 °C)	200.6	199.5	0.0	1.3	162.0	378.5	628.4
5	Total energy use (kWH)	1,557,798.6	2,223,121.8	156.9	497,929.1	885,225.8	1,490,037.0	32,534,133.9

Table 3
Descriptive statistics for building typology. Zeroes added to ambiguous fields, in order to omit these values from the regression models.

	Variable	Mean	Std dev	Min	q25%	Median	q75%	Max
1	Area (m^2)	14,591.0	9563.8	92.9	6271.0	14,492.9	22,900.6	31,122.5
2	Floors	34.4	19.8	1.0	16.0	35.0	53.0	62.0
3	Volume (000 m^2)	2155.2	2070.4	0.4	428.2	1491.8	3291.8	7370.5
4	Buildings on lot	3.8	3.6	1.0	2.0	2.0	2.0	16.0
5	Apartment units	165.4	201.6	0.0	52.0	102.0	203.0	1744.0
6	Commercial spaces	2.6	9.1	0.0	0.0	1.0	3.0	155.0
7	Building age	1189.8	956.1	0.0	0.0	1926.0	1961.0	2009.0
8	Boiler age	15.1	15.6	0.0	0.0	11.0	26.0	64.0
9	Burner age	10.8	13.4	0.0	0.0	5.0	19.0	68.0

It does not use any estimates or assumptions about market share, seasonal or annual energy consumption. Unlike more detailed engineering studies, however, this dataset does not include any detailed energy efficiency measure data. As noted above, if binary data does not show statistically significant effects in initial regression models, then more detailed categorical data is also unlikely to show statistically significant effects either.

Tables 2–4 shows the basic descriptive statistics of the dataset, including the number of buildings, the number of bills per building, and the distribution of the various binary system variables. Numerical data, such as weather and building characteristics (e.g. building area in square meters) in Tables 2 and 3, was center-normalized by subtracting the mean and dividing by the standard deviation in order to ease interpretation of the models, so the relative magnitude of each estimated coefficient could be compared as the relative change in the outcome data as a result of a one standard deviation change in each of the predictors.

All of the criteria for good Bayesian estimation were satisfied. In the model estimation process, flat Bayesian priors were specified for all of the parameters with initial standard deviations of 100,000 were specified in order to allow them to fully explore the parameter space before converging on the point estimates. Monte Carlo Markov Chain (MCMC) methods were used with 5000 iterations, random parameter starting points, 4 separate chains, and the model results were adjusted for burn-in by throwing away the first half (2500) simulated values. Convergence in the chains for all of the parameters was extremely good, with $\hat{R}_{\text{eff}}$ values all less than 1.001, where the criteria for acceptability is less than 1.2 (Gelman and Rubin, 1992).

6. Discussion

Table 5 and Figs. 1–3 present the three different models, all applied to the same observed building energy performance. Model 1 includes all of the available billing and building information, such as could be run with benchmarking and audit data, as will be collected by the Cities of New York and San Francisco. Model 2 is fit only using weather-normalized monthly billing data, similar to the more limited dataset which is provided by the benchmarking law in the City of Seattle. Model 3 is fit using only the building characteristics which could be obtained from an engineering audit,

Table 4
Descriptive statistics for building systems: binary variables.

	Variable	N	Y
1	Direct metered	118	237
2	Comm. spaces with own meter	198	157
3	Comm. spaces with same heating	228	127
4	Comm. spaces with distributed hot water	258	97
5	Comm. spaces with same cooling	312	43
6	Multiple buildings on single boiler	295	60
7	Steam distribution	237	118
8	Radiators	226	129
9	Indicate make control system	69	286
10	Indicate heating distribution problems	32	323
11	Indicate heating problems	326	29
12	Indicate air conditioning equipment	282	73

but without including any historical energy consumption data, which I refer to as the "audit-only" model. These models will be discussed in terms of their overall goodness of fit and the statistical significance of the parameters, and then will be interpreted in terms of practical significance for interpretation.

Evaluating the models in terms of overall goodness of fit or R^2, Model 1 and Model 2 have very high levels of explained variance, 0.8271 and 0.8268, respectively. Between the models there is almost no difference at all in overall model goodness of fit, to at least three significant figures. This implies that Model 2, with only benchmarking data, explains building energy performance just as well as Model 1, which includes engineering audit information of building and system characteristics. When we fit Model 3 with only building and system characteristics, as would be obtained in an engineering audit, we find that while we can obtain a relatively high proportion of explained variance (0.5270), this is still much lower than we obtained before. In comparison, the models for EUI used by the EPA's Portfolio Manager and Energy Star ratings system only have an R^2 of approximately 0.33 (U.S. Environmental Protection Agency, 2011).

In summary, benchmarking data alone, as required by law in the City of Seattle, does just as good a job of explaining the variation in the energy consumption observed in buildings, with or without the additional engineering audit data, as required by the ordinances of the City of New York or San Francisco. Put another way, if you were given only three years of weather-normalized

Table 5
Model results. Means indicate the mean estimates from Bayesian simulations, and parentheses indicate the 95% confidence intervals for each estimate.

Data type	Policy equivalent	New York, San Francisco Model 1: full-information		Seattle Model 2: benchmark only		Model 3: engineering only	
	Variable	Mean	(2.5%, 97.5%)	Mean	(2.5%, 97.5%)	Mean	(2.5%, 97.5%)
Monthly bill	Precipitation (mm)	−0.035	(−0.051, −0.019)	−0.035	(−0.052, −0.018)	−0.028	(−0.052, −0.003)
	HDD (18 °C)	−0.052	(−0.084, −0.017)	−0.052	(−0.085, −0.016)	−0.057	(−0.091, −0.023)
	CDD (18 °C)	0.122	(0.085, 0.158)	0.121	(0.087, 0.159)	0.122	(0.088, 0.155)
	Year of bill	0.114	(0.099, 0.129)	0.114	(0.099, 0.129)	0.106	(0.082, 0.131)
	Building age	−0.202	(−0.355, −0.055)			−0.171	(−0.197, −0.145)
Building	Area (m^2)	−1.204	(−1.509, −0.897)			−1.412	(−1.465, −1.36)
	Floors	−0.004	(−0.26, 0.253)			−0.049	(−0.094, −0.007)
	Volume (m^3)	0.133	(−0.253, 0.515)			0.232	(0.169, 0.297)
	Buildings in complex	−0.206	(−0.373, −0.042)			−0.033	(−0.065, −0.003)
	No. apartment units	0.37	(0.198, 0.542)			0.362	(0.333, 0.391)
	No. comm. spaces	0.046	(−0.109, 0.206)			0.098	(0.072, 0.124)
	Boiler age	0.004	(−0.186, 0.187)			−0.007	(−0.038, 0.024)
	Burner age	−0.047	(−0.232, 0.134)			−0.099	(−0.129, −0.068)
Comm. spaces	Separate meter?	0.34	(−0.073, 0.745)			0.005	(−0.107, 0.113)
	Same heating?	0.102	(−0.46, 0.678)			0.351	(0.236, 0.472)
	Same DHW?	−0.181	(−0.66, 0.295)			−0.064	(−0.145, 0.017)
	Same cooling?	−0.433	(−0.932, 0.063)			−0.471	(−0.558, −0.386)
Systems	Direct-meter?	0.086	(−0.3, 0.467)			0.139	(0.075, 0.201)
	Mult. bldg. on boiler?	−0.088	(−0.48, 0.302)			−0.449	(−0.527, −0.371)
	Steam dist.?	−0.174	(−0.549, 0.197)			−0.304	(−0.368, −0.24)
	Radiators?	0.524	(0.112, 0.933)			0.661	(0.591, 0.731)
	Make ctrl system?	0.022	(−0.441, 0.491)			0.147	(0.07, 0.225)
	Heating dist. prob.?	0.143	(−0.518, 0.824)			0.146	(0.033, 0.262)
	Heating prob.?	0.295	(−0.194, 0.793)			0.222	(0.135, 0.306)
	AC Units?	−0.093	(−0.537, 0.371)			−0.104	(−0.184, −0.024)
Variance	Monthly	0.046	(0.019, 0.088)	0.045	(0.019, 0.084)		
	Building	1.144	(1.04, 1.257)	1.75	(1.604, 1.917)		
	Residual errors	0.787	(0.776, 0.797)	0.787	(0.777, 0.798)		
Model fit	Expl. variance (R^2)	0.8271		0.8268		0.5270	

energy bills, you would still be able to obtain an equally good prediction of the energy use of a given building in the next year, without needing to know anything else about the building characteristics or engineering systems. This clearly shows that auditing information is less valuable than benchmarking information in predicting the energy performance of a given building.

Figs. 1–3 visualizes and explains further these model results in terms of the statistical significance of the individual predictive variables. The horizontal, logarithmic axis is in terms of orders of magnitude of the energy use intensity, so estimates for each of the coefficients are identified with a dot, while the 95% confidence intervals and the standard deviation intervals are indicated with thick and thin line segments, respectively. The null hypothesis is that each of the coefficients has no effect on the observed outcome of total building energy consumption; this is only true if the thick and thin line segments cross the zero axis indicated by the vertical dotted line.

In Fig. 1, one can see that very few of the building and system characteristics are statistically significant, since in many cases the 95% confidence intervals and the standard deviation intervals cross the zero (center) line. In addition, because the numerical variables have been center-normalized, we can directly compare the effect of a standard deviation change in any of the predictors with the variation between buildings. The magnitude, or effect sizes, of all of the system characteristics are considerably smaller than the effect of building-level effects, in most cases by an order of magnitude, since the horizontal axis is a logarithmic scale. Square footage of the building still has a significant impact on the EUI of the building, meaning that total energy consumption goes down with the effect of area squared, though that there is an effect of area on energy consumption is not surprising. The presence of radiators is the only building system that has a significant impact on energy consumption.

Fig. 2 shows the same results of Fig. 1, since each of the climate variables is significantly smaller than the effect of building-level variation. Fig. 3 shows a slightly different though not inconsistent result: although each of the variables is statistically significant in the audit model, they are only significant in the absence of the benchmarking data, and the overall model fit is worse, as mentioned above.

Taken together, the model results show that, surprisingly, almost none of the individual building-level predictors from engineering audits are statistically significant when compared to the data that could be obtained by regular benchmarking. In contrast, building-specific effects in observed energy consumption are persistent year over year, relatively more important in their magnitude, and statistically significant. This states in clear statistical terms that benchmarking information is relatively more important than audit information when attempting to compare the relative energy performance of two buildings. Policymakers should therefore consider benchmarking to be a low-cost but cost-effective information disclosure policy when compared to more-expensive audits that do not add any additional prediction value about the expected energy performance for a given building.

Furthermore, since the sample of buildings has previously been found to be broadly representative of the overall population, the relative magnitude of the effect sizes are unlikely to change even if many more audits are conducted. If sample size increases, although the errors in the estimates will decrease as in Model 3, the magnitude or effect sizes for particular building characteristics

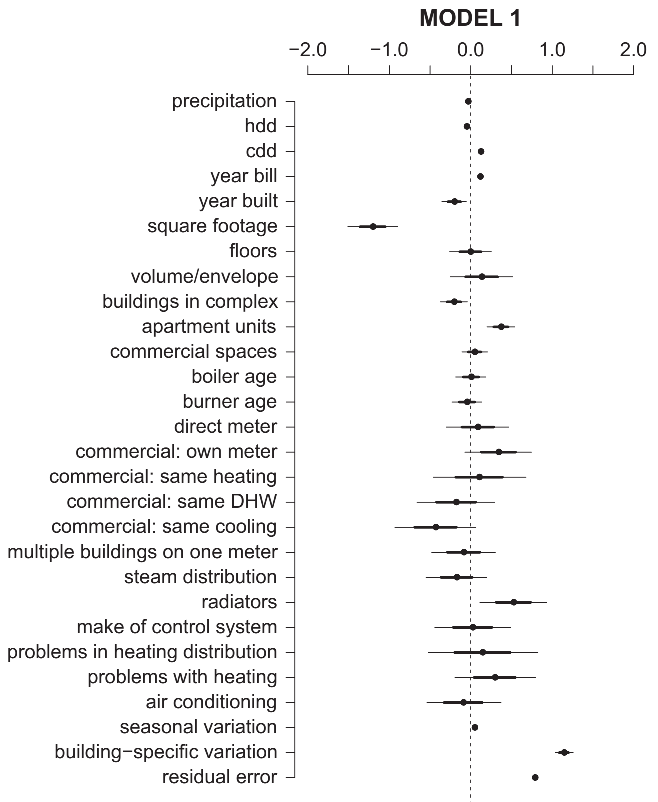

Fig. 1. Estimation results for model. Center dots indicate the mean estimate, the thicker bars indicate the standard error of the estimate, and the thinnest bars indicate the 95% confidence interval. As the graph shows, the climate variables are significant but not highly influential (with small magnitude). The square footage of the building appears to be highly influential and statistically significant (because the error bars do not include zero). Most surprisingly, for most of the building and system level variables, they are neither influential nor statistically significant. By far, the most important finding is that the building-level variation is significantly higher than the seasonal variation, or any other building system or characteristic.

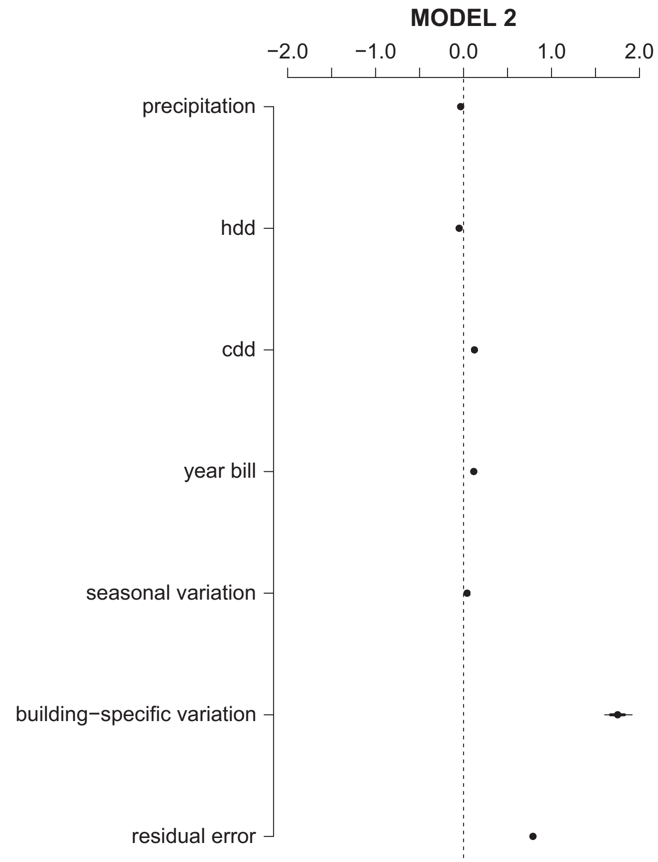

Fig. 2. Estimation results for model. Center dots indicate the mean estimate, the thicker bars indicate the standard error of the estimate, and the thinnest bars indicate the 95% confidence interval.

or systems will still be much less than that of the building-specific effects found in Models 1 and 2. This means that variations from sources specific to the building, such as the unobserved effects of operations and/or occupancy, may be much more important than changes to building systems, or the efficiency measures that have been typically recommended by building engineers and consultants. While this is in keeping with recent research on the effects of occupancy on building energy consumption (Lutzenheiser, 1993; Guerra Santin et al., 2009; Deuble and de Dear, 2012), it is worth noting two things: at present there are no major policy initiatives that require disclosure of occupancy information because of privacy and commercial concerns, and it remains very difficult to measure occupancy information that could be used for benchmarking. Information disclosure policies that report occupancy are purely theoretical at present.

7. Conclusions

This paper began by reviewing the important contribution of buildings to worldwide energy consumption and carbon emissions, and then described the policy context of recent efforts to pass energy disclosure laws that are intended to transform markets for energy efficiency. These policy developments were connected to the overall background of disclosure and information laws used in environmental regulation and have been studied extensively in the literature.

This paper then applied a flexible statistical model form to a comprehensive dataset of building energy consumption and engineering audits, in order to determine which data is most valuable in order to predict the log EUI of individual buildings, and could therefore be expected to transform markets. In almost all cases, the effects of individual building and system characteristics are statistically insignificant relative to the building-level variation, which could be revealed by a well-designed benchmarking program.

What the results also imply is that even controlling for all observed characteristics, prospective buyers and tenants of buildings should pay close attention to the annualized energy consumption of individual buildings, and that policymakers should consider energy efficiency relative to current energy use and not technical systems. These results will hopefully shift attention to capturing operational efficiencies and behavioral changes in building occupants, rather than changing the systems of buildings, as engineering-economic models frequently emphasize. It is relatively cheap to capture operating efficiencies instead of retrofitting building systems.

Considering these results in terms of market transformation, while the growth in new data from benchmarking is intended to reduce information asymmetries, additional information from engineering audits may not allow improved systematic prediction of energy performance. It is possible that there are additional benefits to engineering audits which may not be captured by this study. For example, a building owner who commissions an engineering audit may receive a list of proposed energy conservation measures, and this may spur them to take action on energy efficiency. On the other hand, it is also possible that these energy conservation measures and system changes will not yield the

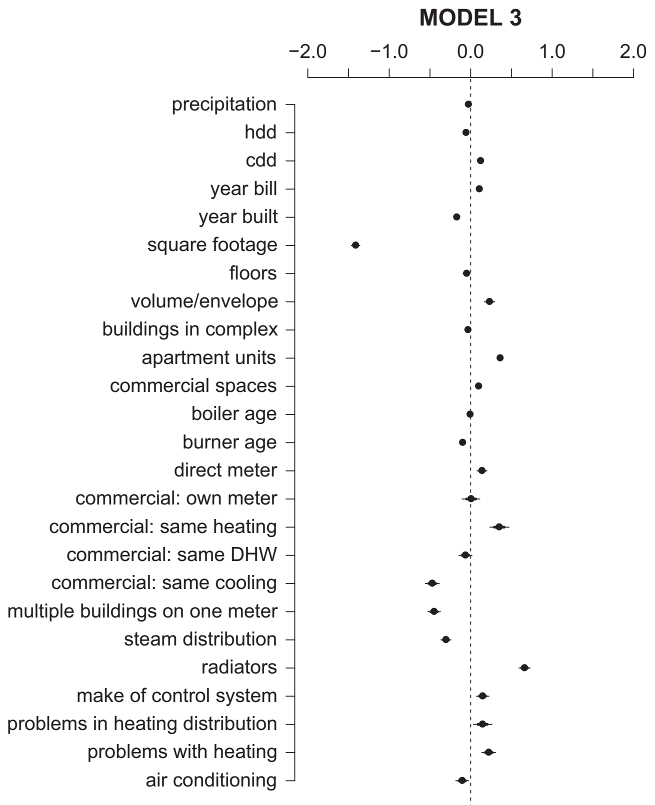

Fig. 3. Estimation results for model. Center dots indicate the mean estimate, the thicker bars indicate the standard error of the estimate, and the thinnest bars indicate the 95% confidence interval.

expected savings, as Shapiro (2011) notes, resulting in a disappointed or demotivated customer for energy efficiency.

These results also lead to interesting potential conclusions about how policymakers, real estate owners, and consultants might communicate when analyzing the overall market for energy efficiency. First, these results imply that there is significant variation in building energy use regardless of systems, meaning that it very possible that identical buildings may have different levels of energy use intensity. Second, if energy use intensity does not directly result from the systems in a building, then it may be much cheaper to focus on operational efficiency improvements rather than technical or system upgrades. Third, if the only variable that truly matters is the observed energy use, rather than their particular systems, then policymakers could simply inform all building owners that they need to meet a uniform standard for energy use intensity when compared to similar buildings, say by facility type. Or, owners of multiple buildings could try to reduce their energy use by an arbitrary standard, say to the median or better, with the expectation that all buildings, on average, have some potential for improvement regardless of systems. This approach would be quite different from the existing utility approach that assesses energy efficiency potential based on the technical and economic feasibility of particular energy conservation measures. Consultants that are confident that operational efficiencies remain un-captured could proposed to reduce energy consumption by a certain amount from the existing baseline, without necessarily having to resort to installing physical conservation measures. However, ESCOs may want additional control or clauses to affect occupant behavior, since this may have a relatively larger effect than anything that they may be able to install.

One significant limitation to this research is that even if this analysis finds that engineering audits do not explain building-level variation, the source of this heterogeneity remains unobserved or unexplained. Other information that might explain the observed energy use would be information about operations, occupancy, and existing conservation measures, but it is very difficult to obtain this comprehensive information about buildings consistently over long periods of time, and for reasonable sample sizes. If this information could be obtained, then it may help to explain unobserved heterogeneity in buildings and point towards further refinement of disclosure policies. However, as noted above, at present there are no major policies being proposed that would mandate systematic measurement of occupancy or energy conservation measures, and therefore these types of data cannot yet be used to predict the energy performance of buildings in a practical manner.

Furthermore, these results should not be taken to imply that auditing information has no additional value in energy efficiency. Whether or not it is actually possible to reduce energy consumption by an arbitrary percentage may differ from building to building, particularly if operational energy efficiency measures have already been implemented. Audits may still provide important information to owners and occupiers about the energy efficiency measures that could be further undertaken to reduce energy use. However, these results can be read as implying that auditing information may not provide any additional predictive power beyond benchmarking, meaning that a prospective buyer or tenant may have no use for this additional and costly level of information to predict the future energy use of a building. If policymakers pass disclosure and auditing requirements, then they should be connected either to different theories of market transformation, or to different stages of energy efficiency efforts. Finally, these results imply that we should perhaps be looking for structural incentives, such as fuel prices, carbon taxes, or other such signals, that would encourage all owners to reduce their current levels of energy use regardless of their current building characteristics.

Acknowledgments

FS Energy, an energy advisory subsidiary of First Service Residential, provided utility billing and building typology data, as well as funding to assemble and clean the data. Analysis of the data was sponsored thanks to the U.S. Department of Energy, Energy Efficient Buildings HUB, Grant number DE-EE0004261. Aaron Mehta helped greatly with the cleaning and assembly of the dataset. The research question, design, writing of the report, and conclusions were solely decided by the author.

Many thanks to Clinton Andrews, John Landis, Loren Lutzenheiser, Laurie Kerr, and Ron Merhige for their insights and feedback.

References

Akerlof, G.A., 1970. The market for "lemons": quality uncertainty and the market mechanism. The Quarterly Journal of Economics 1, 488–500.

Allcott, H., Greenstone, M., 2012. Is there an energy efficiency gap? Journal of Economic Perspectives 26 (1), 3–28.

Anderson, S.T., Newell, R.G., 2004. Information programs for technology adoption: the case of energy-efficiency audits. Resource and Energy Economics 26 (March (1)), 27–50.

ASHRAE, June 2011. Procedures for Commercial Building Energy Audits, 2nd ed. American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Bae, H., Wilcoxen, P., Popp, D., 2010. Information disclosure policy: do state data processing efforts help more than the information disclosure itself?. Journal of Policy Analysis and Management 29 (December (1)), 163–182.

Banerjee, A., Solomon, B., 2003. Eco-labeling for energy efficiency and sustainability: a meta-evaluation of US programs. Energy Policy 31 (2), 109–123.

Belzer, D., April 2009. Energy Efficiency Potential in Existing Commercial Buildings: Review of Selected Recent Studies. Technical Report PNNL-18337, Pacific Northwest National Laboratory.

Blumstein, C., Krieg, B., Schipper, L., York, C., 1980. Overcoming social and institutional barriers to energy conservation. Energy 5 (4), 355–372.

Burr, A., Keicher, C., Leipziger, D., July 2011. Building Energy Transparency: A Framework for Implementing U.S. Commercial Energy Rating and Disclosure Policy. Technical Report, Institute for Market Transformation.

California Energy Commission, January 2000. How to Hire an Energy Auditor to Identify Energy Efficiency Projects. Technical Report, 400-00-001C, California Energy Commission.

Chung, W., 2011. Review of building energy-use performance benchmarking methodologies. Applied Energy 88 (May (5)), 1470–1479.

Delmas, M., Montes-Sancho, M.J., Shimshack, J.P., 2010. Information disclosure policies: evidence from the electricity industry. Economic Inquiry 48 (2), 483–498.

Dempster, A.P., Rubin, D.B., Tsutakawa, R.K., 1981. Estimation in covariance components models. Journal of the American Statistical Association 76 (374), 341–353.

Deuble, M.P., de Dear, R.J., 2012. Green occupants for green buildings: the missing link?. Building and Environment 56 (October (0)), 21–27.

Eddy, W.F., Marton, K. (Eds.), 2012. Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys. National Academies Press.

Efron, B., Morris, C., 1975. Data analysis using stein's estimator and its generalizations. Journal of the American Statistical Association 70 (350), 311–319.

Eichholtz., Piet., Nils Kok., John M. Quigley., 2010. Doing well by doing good? Green office buildings. The American Economic Review 100 (December (5)), 2492–2509. http://dx.doi.org/10.2307/41038771.

Faraway, J.J., 2005. Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models, vol. 66. Chapman & Hall/CRC.

Florance, A., Miller, N., Peng, R., Spivey, J., 2010. Slicing, dicing, and scoping the size of the U.S. commercial real estate market. Journal of Real Estate Portfolio Management 16 (January (2)), 101–118.

Fuerst, F., McAllister, P., 2011. Green noise or green value? Measuring the effects of environmental certification on office values. Real Estate Economics.

Fung, A., 2007. Full Disclosure: The Perils and Promise of Transparency. Cambridge University Press, Cambridge, New York.

Gelfand, A.E., Smith, A.F., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85 (410), 398–409.

Gelman, A., 2005. Analysis of variance – why it is more important than ever. The Annals of Statistics 33 (1), 1–53.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7 (November (4)), 457–472, ArticleType: research-article / Full publication date: Nov., 1992 / Copyright 1992 Institute of Mathematical Statistics.

Guerra Santin, O., Itard, L., Visscher, H., 2009. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy and Buildings 41 (November (11)), 1223–1232.

Hewitt, D., Frankel, M., Cohan, D., 2010. The future of energy codes. In: Conference Proceedings of the 2010 ACEEE Summer Study on Energy Efficiency in Buildings. Asilomar, CA.

Hirst, E., Brown, M., 1990. Closing the efficiency gap: barriers to the efficient use of energy. Resources, Conservation and Recycling 3 (June (4)), 267–281.

Howard, B., Parshall, L., Thompson, J., Hammer, S., Dickinson, J., Modi, V., 2012. Spatial distribution of urban building energy consumption by end use. Energy and Buildings 45 (February (0)), 141–151.

Ingle, A., Moezzi, M.M., Lutzenheiser, L., Hath, Z., Lutzenhiser, S., Van Clock, J., Peters, J., Smith, R., Heslam, D., Diamond, R., May 2012. Behavioral Perspectives on Home Energy Audits: The Role of Auditors, Labels, Reports, and Audit Tools on Homeowner Decision-Making. Technical Report.

Itron, Inc., 2006. Evaluation of the School Energy Efficiency Program. Technical Report CALMAC DRI0002.01; 1190-04.

Jaffe, A.B., Stavins, R.N., 1994. Energy-efficiency gap what does it mean? Energy Policy 22 (10), 804–810.

Judkoff, R., Neymark, J., 2006. Model validation and testing: the methodological foundation of ASHRAE standard 140. Transactions - American Society of Heating, Refrigerating and Air Conditioning Engineers 112 (2), 367.

KEMA, Inc., 2006. Evaluation of 2004–2005 Designed for Comfort: Efficient Affordable Housing Program. Technical Report CALMAC HMG0002.01.

Kerr, L., February 2013. Personal communication.

Kerr, L., Beber, H., Hope, D., June 2012. New York City Local Law 84 Benchmarking Report. Technical Report, City of New York Office of Long-Term Planning and Sustainability.

Konar, S., Cohen, M.A., 1997. Information as regulation: the effect of community right to know laws on toxic emissions. Journal of Environmental Economics and Management 32 (January (1)), 109–124.

Krarti, M., January 2011. Energy Audit of Building Systems: An Engineering Approach, 2nd ed. Taylor & Francis US.

Levine, M.D., Koomey, J.G., McMahon, J.E., Sanstad, A.H., 1995. Energy efficiency policy and market failures. Annual Review of Energy and the Environment 20 (1), 535–555.

Lindley, D.V., Smith, A.F., 1972. Bayes estimates for the linear model. Journal of the Royal Statistical Society. Series B (Methodological), 1–41.

Lutzenheiser, L., 1993. Social and behavioral aspects of energy use. Annual Review of Energy and the Environment 18 (November), 247–289.

Massey, R.I., 2011. Program assessment at the 20 year mark: experiences of Massachusetts companies and communities with the toxics use reduction act (TURA) program. Journal of Cleaner Production 19 (March (5)), 505–516.

Mazria, E., 2003. It's the architecture, stupid!. WORLD AND I 18 (10), 138–145.

McKinsey, July 2009. Unlocking Energy Efficiency in the U.S. Economy. Technical Report, McKinsey & Company.

Mehta, A., August 2013. Fs energy data integrity process, Personal communication.

Miller, N., Spivey, J., Florance, A., 2008. Does green pay off? Journal of Real Estate Portfolio Management 14 (4), 385–400.

Mills, E., 2011. Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the United States. Energy Efficiency 4 (May (2)), 145–173.

Norford, L., Socolow, R., Hsieh, E., Spadaro, G., 1994. Two-to-one discrepancy between measured and predicted performance of a 'low-energy' office building: insights from a reconciliation based on the DOE-2 model. Energy and Buildings 21 (2), 121–131.

Pacala, S., Socolow, R., 2004. Stabilization wedges: solving the climate problem for the next 50 years with current technologies. Science 305 (5686), 968.

Palmer, K., Walls, M., Gordon, H., Gerarden, T., January 2012. Assessing the Energy Efficiency Information Gap: Results from a Survey of Home Energy Auditors. Technical Report RFF DP 11-42, Resources for the Future.

Pang, X., Wetter, M., Bhattacharya, P., Haves, P., 2012. A framework for simulation-based real-time whole building performance assessment. Building and Environment 54 (August (0)), 100–108.

Patten, D.M., 1998. The impact of the EPA's TRI disclosure program on state environmental and natural resource expenditures. Journal of Accounting and Public Policy 17 (4–5), 367–382.

Pauler, D.K., Wakefield, J.C., Kass, R.E., 1999. Bayes factors and approximations for variance component models. Journal of the American Statistical Association 94 (December (448)), 1242–1253.

Pérez-Lombard, L., Ortiz, J., Pout, C., 2008. A review on buildings energy consumption information. Energy and Buildings 40 (3), 394–398.

Plummer, M., 2012. JAGS.

Plummer, M., Feb. 2013. RJAGS: Bayesian Graphical Models Using MCMC.

R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2008.

Robert Mowris and Associates, 2007. Evaluation, Measurement, and Verification Report for the Small Nonresidential Energy Fitness Program #1409-04. Technical Report CALMAC RHA0002.01.

Robert Mowris and Associates, 2008. Evaluation, Measurement, and Verification Report for the Moderate Income Comprehensive Attic Program #1082-04. Technical Report CALMAC BOE0001.01; 1082-04.

Rosenthal, E., January 2012. Hard Truths about Disclosure. The New York Times.

Ryan, E.M., Sanquist, T.F., 2012. Validation of building energy modeling tools under idealized and realistic conditions. Energy and Buildings 47 (April (0)), 375–382.

Schleich, J., 2004. Do energy audits help reduce barriers to energy efficiency? An empirical analysis for Germany. International Journal of Energy Technology and Policy 2 (January (3)), 226–239.

Shapiro, I., 2011. 10 common problems in energy audits. ASHRAE Journal (February), 26–32.

Stavins, R., 2003. Experience with market-based environmental policy instruments. Handbook of Environmental Economics 1, 355–435.

Turner, C., Frankel, M., March 2008. Energy Performance of LEED for New Construction Buildings. Technical Report, New Buildings Institute, Vancouver, WA.

Ürge-Vorsatz, D., Novikova, A., 2008. Potentials and costs of carbon dioxide mitigation in the world's buildings. Energy Policy 36 (February(2)), 642–661.

U.S. Department of Energy, September 2011. A Guide to Energy Audits. Technical Report, Richland, WA.

U.S. Environmental Protection Agency, March 2011. Energy Star Performance Ratings: Technical Methodology. Technical Report, U.S. Environmental Protection Agency.

USEIA, September 2012. Annual Energy Review. Technical Report, U.S. Energy Information Administration.

Weil, D., Fung, A., Graham, M., Fagotto, E., 2006. The effectiveness of regulatory disclosure policies. Journal of Policy Analysis and Management 25 (December (1)), 155–181.

Winston, C., 2008. The efficacy of information policy: a review of Archon Fung, Mary Graham, and David Weil's full disclosure: the perils and promise of transparency. Journal of Economic Literature 46 (3), 704–717.

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Zerbe, R.O., Dively, D.D., 1994. Benefit-Cost Analysis: In Theory and Practice. Harper-Collins.