Measuring Energy Savings from Benchmarking Policies in New York City

Ting Meng, Massachusetts Institute of Technology
David Hsu, Massachusetts Institute of Technology
Albert Han, University of Pennsylvania

ABSTRACT

Benchmarking policies, also known as disclosure ordinances, are being pursued in many countries and at all levels of government. These policies are intended to transform the market for energy efficient buildings by requiring owners to measure and disclose their energy use, therefore making energy efficiency investments more visible and therefore valuable. Two critical questions to improving these policy efforts are: first, how much energy do these policies save? and second, what particular aspects of these policies are most effective? To answer these questions, this study exploits how different aspects of these policies were phased-in to different groups of buildings over the first four years of the City of New York's benchmarking ordinance. By identifying treatment and control groups within each stage of implementation, and then applying a novel difference-in-differences strategy, we can causally attribute observed declines in energy consumption to specific owner behaviors and policy mechanisms. Our analysis indicates that in comparison with the control group and before the policies were implemented in 2011, the total disclosure policy can be credited with a 6% reduction in building energy use intensity (EUI) three years later and a 14% reduction in EUI four years later; the disclosure of Energy Star scores decreased building EUI by 9% three years later and 13% four years later. The two sets of independent findings are a consequence of the policy design and different control groups.

Introduction

Many governments, worldwide and at all levels, have already implemented benchmarking or energy disclosure ordinances for buildings in order to transform the market for energy efficient buildings (Burr et al. 2011; Hinge and Miclea 2015). These policies require building owners to gather and disclose their energy use, sometimes on a public website, which is intended to lead to a virtuous cycle: injecting this information into the market then makes it possible for potential buyers, renters, and tenants to see which buildings are more energy efficient, therefore valuing this in their purchase or leasing decisions, and therefore giving owners additional incentives to invest in energy efficiency, and so on.

With more cities considering similar policies or ordinances (NRDC and IMT 2015), two critical questions to improving these policy efforts are: first, how much energy do these policies save? and second, what particular aspects of these policies are most effective? Since it has also been more than three years since the earliest cities in the United States implemented their benchmarking policies, it should be possible to measure the short- and medium-term effects of these policies, although it is too early to assess the magnitude or persistence of long-term effects (DOE 2015).

To answer these questions, this study exploits how different aspects of these policies were phased-in to different groups of buildings over the first four years of the City of New York's benchmarking ordinance (the city and its government will both be henceforth referred as "NYC"). By creating a balanced panel dataset, identifying treatment and control groups within each stage of implementation, and then applying a novel implementation of a difference-in-differences strategy, we can causally attribute observed declines in energy consumption to specific owner behaviors and policy mechanisms.

Few studies have previously focused on analyzing and evaluating these policies. This study contributes to the existing literature on the evaluation of benchmarking policies in three ways: first, this study investigates the underlying causal effects, rather than correlations, between disclosure policies and energy use intensity (EUI); second, beyond evaluating the overall disclosure effect, to our knowledge this is the first study to further investigate the disclosure effects of particular policy aspects (i.e., disclosure of Energy Star scores); and third, this method could be generalized easily to other cities, so long as their policies are implemented in a phased process. All of these contributions should be useful to the design and implementation of benchmarking policies in other cities and in the future.

This study is organized into the following sections. The next (second) section introduces a broad framework of various energy policies targeting the building sector, with a particular focus on studies that evaluate benchmarking policies. The third section presents a conceptual framework of how benchmarking policies are expected to reduce the energy use of buildings as well as the casual inference study design and the associated empirical model. Next, the fourth section summarizes the data source, descriptive statistics, and estimation results. Finally, the fifth section presents the conclusions of the study, the limitations of the data and results, and possible implications for policy design and implementation.

Our analysis indicates that in comparison with the control group and before the policies were implemented in 2011, the total disclosure policy can be credited with a 6% reduction in building energy use intensity (EUI) three years later and a 14% reduction in EUI four years later; the disclosure of Energy Star scores decreased building EUI by 9% three years later and 13% four years later. The two sets of independent findings are a consequence of the policy design and different control groups.

Related Work

A growing number of studies describe a wide range of new policies designed to reduce energy use of buildings. These policies include energy certification, technical standards, and the mandatory disclosure of energy information (Brown et al. 2002; Beerepoot and Beerepoot 2007; Perez-Lombard et al. 2008; Andaloro et al. 2010).

Among various energy policies, benchmarking policies have become "an important frontier of government innovation" and "one of the most promising public policy tools" to improve building energy performance and efficiency (Kontokosta 2013; Weil et al. 2006). In these benchmarking policies, building owners are required to measure and disclose their energy use information through the U.S. Environmental Protection Agency's (EPA) free online benchmarking tool called Portfolio Manager. EPA provides benchmarked buildings with their Energy Star scores (for eligible properties) based on their peer buildings' performance. These policies aim to give building owners and potential buyers a better understanding of buildings' energy use performance, and eventually shifting the market towards increasingly efficient and high-performing buildings.

NYC was one of the earliest cities to implement a benchmarking policy, introducing Local Law 84 in 2009, which required commercial and multifamily buildings above 50,000 square feet to disclose their energy use data. Compliance was high from the beginning, starting above 75% of eligible buildings, resulting in reporting of energy use in more than 10,000 buildings.

In spite of the growing popularity of benchmarking policies, with more cities considering implementing them, only a few studies have explored this particular type of policy and evaluated its effectiveness. EPA reported that all Portfolio Manager benchmarked buildings saved 7% of energy over the period from 2008 to 2011, and had a 6-point-increase of Energy Star scores (EPA 2012). Similarly, in NYC's latest benchmarking report (issued in September 2014), the median EUI of benchmarked office and multifamily buildings went down 13% and 12% from 2011 to 2012, respectively (NYC 2014).

However, as Palmer and Walls (2015) point out, these conclusions are based on average energy use over time, meaning that these are correlations and not causally attributable results. Energy use in buildings could be affected by confounding factors such as broad trends in the economy, energy sector, or real estate sector, and therefore not necessarily benchmarking policies. In order to attribute the effect to benchmarking policies, Palmer and Walls (2015) instead use National Council of Real Estate Investment Fiduciaries (NCREIF) data and also apply a difference-in-differences regression model to develop a causal attribution of energy savings to benchmarking policies, finding approximately a 3% reduction in utility costs after the passage of benchmarking policies. While this is an excellent precedent for our subsequent work, one of the limitations in their study is that they can only investigate utility expenditure per square foot in a limited set of investor-owned office and retail buildings. Since benchmarking policies are intended to affect energy consumption within the broader population of buildings, our approach is targeted towards evaluating the effect of benchmarking policies on the overall population, and can be integrated into future implementations of benchmarking policy.

Conceptual Framework, Empirical Model, and Research Design

In this section we introduce the following: first, a conceptual framework for the ways in which benchmarking policy might act to reduce energy consumption; second, the difference-in-difference model for causal attribution of policy outcomes to implemented policies; and third, our research design which exploits features of the phasing-in of NYC's benchmarking ordinance to different groups of buildings.

Benchmarking or disclosure policies gather and provide information to the market. There are three different possible mechanisms by which making owners more aware of their energy use may influence building owners to use less energy in subsequent years, but each of these mechanisms work in different ways:

- preparing energy use information (N) – the act of gathering energy information may draw building owner's attention to their energy use, and this awareness may result in reductions in energy use;
- disclosing basic energy information (I) to the public – may affect the decisions of potential tenants and buyers, further motivating owners to reduce energy use; or
- disclosing Energy Star scores to the public (S) – works similar to energy information above, but using an industry standard to communicate to the market.

Figure 1 displays three main pathways by which benchmarking policies may affect building energy consumption.

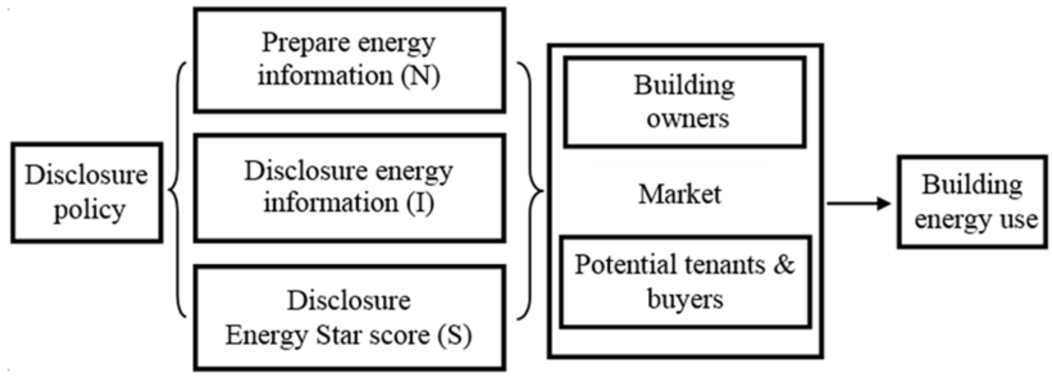

Figure 1: Conceptual Framework

We then use the difference-in-differences (DinD) model to establish a causal and measurable link between observed outcomes and implemented policies. This model is often used for policy or program evaluation for a wide range of disciplines and purposes (Duflo et al. 2007; Palmer and Walls 2015). The strong assumption in this model is that the treatment and control groups would have followed the same trajectory in the absence of the treatment, allowing the effect of the treatment to be calculated as the difference of the changes in the treatment and control groups over time. Figure 2 shows the basic set-up:

	Before	After
Treatment	0	1
Control	0	0

Figure 2: The difference-in-differences (DinD) model

In the typical difference-in-differences model shown here, "1" indicates the observations affected by the new implementation of certain policies or programs (the treatment), while "0" indicates that observations are not influenced by the new policy implementation (before for the treatment, or before and after for the control group). The before-after difference in the treatment group represents the sum of the policy and confounding factors, and the before-after difference in the control group represents only the confounding factors. Therefore, the difference of these before-after differences between treatment and control groups (the second difference) represents the isolated policy effect.

By identifying treatment and control groups both before and after policies take effect, and measuring policy outcomes of interest for each group and time, we can then use the difference-in-differences casual study design to estimate effects using an econometric regression approach. The equation specification for the DinD model is:

$$Y = \beta_0 + \beta_1 A + \beta_2 T + \beta_3 A*T + X\beta + \varepsilon$$

where Y denotes the outcome of interest; A is an indicator variable for periods after the new policy has been implemented; T is an indicator variable for treatment groups; A*T, or A times T, is the interaction term between A and T; X is a vector of building features; β's are corresponding coefficients; and ε is the stochastic error term. β_3, the parameter of the interaction term A*T, measures the isolated policy effect.

Our research design therefore exploits the fact that aspects of NYC's benchmarking policy were implemented in a series of different steps and phases. In 2011, NYC first disclosed energy data for its publicly-owned buildings of all types. In 2012, NYC first required disclosure of energy data for privately-owned buildings, including both commercial and multifamily buildings. Also, at this point Energy Star scores existed for some classes of commercial properties but not multifamily buildings; EPA did not launch Energy Star scores for multifamily buildings until 2014. Our strategy was to divide the buildings into different treatment and control groups, based on whether buildings were supposed to be affected by the new disclosure policy and whether Energy Star scores were disclosed along with the other basic energy use information. We conducted analyses comparing the treatment and control groups for similar facility types, to verify that any observed effects were not the result of comparing dissimilar buildings. To test the effect of benchmarking and disclosure, we compared only the same facility type – offices – since NYC does own some commercial office buildings that are similar to those in the private sector.

One novel aspect of our use of the DinD model is that the control group can be either a group which the policy does not affect, using indicator variables 0 for before and 0 for after, as in Figure 1; or the indicator variables can be 1 for before and 1 for after; the mathematics of the DinD estimation works out the same. We therefore use privately-owned office buildings that are affected by the newly implemented benchmarking policies as the treatment group. Also, in one model we use publicly-owned office buildings that have been consistently benchmarked as a control group, and in a second model we use privately-owned office buildings that have consistently not received Energy Star scores as a control group.

Furthermore, in the difference-in-differences model, first- and second-differences between the equations are then calculated in order to untangle multiple aspects of the policy if they were introduced in different periods. Additional control factors influencing building energy use are also included for each comparison between the before and after periods. We then use comparisons between two different treatment and control groups to isolate the effects of total disclosure effect (I+S) and as well as disclosure of Energy Star scores (S) described above.

The two tables below show how the study design is used to isolate key effects (I+S, S) through two key comparisons. In Table 1, the treatment group was privately-owned scored offices that were newly benchmarked and reporting Energy Star scores (T, n=261), and the control group was publicly-owned scored offices that have been benchmarked before and also received Energy Star scores (C1, n=75). Also in Table 2, to analyze the effect of disclosing Energy Star scores, the same treatment group as above is compared to privately-owned non-scored buildings that have between 20% and 100% office space (C2, n=39), but with no Energy Star scores for various reasons. For example, buildings with less than 50% office space do not receive scores because they are classified as multifamily buildings, and Energy Star scores did not exist at the time for these facility types.

Table 1: Study design of quantifying the total benchmarking effects

	Before	After	
Privately-owned, scored (T, n=261)	N	N+I+S+O	I+S+ O (1st diff., before-after)
Publicly-owned, scored (C1, n=75)	N+I+S	N+I+S+O	O (1st diff., before-after)
			I+S (2nd diff., treatment & control)

Table 2: Study design of quantifying the effect of disclosing Energy Star scores

	Before	After	
Private, scored (T, n=261)	N	N+I+S+O	I+S+ O (1st diff., before-after)
Private, non-scored (C2, n=39)	N	N+I+ O	I+ O (1st diff., before-after)
			S (2nd diff., treatment & control)

Letters denote mechanisms working in each period on each group. N: preparing energy use information; I: the disclosure effect of basic energy use information; S: the disclosure effect of Energy Star scores; O: other co-founded factors.

Data

We assembled a balanced panel dataset by joining two sets of data – Portfolio Manager (PM) data for NYC buildings, obtained from the NYC Mayor's Office of Long-Term Planning and Sustainability; and the City of New York's Primary Land Use Tax Lot Output (PLUTO). It covers a four-year time period from 2011 to 2014, where the reported year indicates data gathered in the year before. This panel data set provides energy use and efficiency data for both privately-owned and publicly-owned buildings. Specifically, it includes site and source energy use intensity, building use by types (i.e., electricity, natural gas, fuel oil, and street steam), and as well as Energy Star scores if it is applied. Other important building characteristics such as building and lot sizes, zoning, and built year are also available in the PLUTO data.

A series of data cleaning steps were carried out. First, observations with duplicated records or missing essential variables such as weather normalized source energy use intensity and building area were deleted. Second, buildings with duplicated identification numbers were removed in order to select those that have one building per each lot. Third, buildings reporting extremely high or low energy use efficiency (top 1% and bottom 1% of weather normalized source energy use intensity) and buildings with unreasonable energy use percentages (100% fuel use or zero electricity use) were eliminated as outliers. Fourth, we only kept buildings with four years of reporting in order to create a balanced panel to maintain consistent policy evaluation over time (DOE 2015). Because this study tracks the same buildings across years, standard errors in the estimation are clustered by the building identification numbers to deal with the correlations in energy use over years for each building. Fifth, we dropped the variable floor area ratio (FAR) from the model specification due to its data quality, and also because it is highly correlated with the number of floors that has been included in the model.

The dependent variable of interest is the logarithm of weather-normalized source energy use intensity (EUI). The logarithmic transform was employed to compensate for observed left hand skewness in the data. This also enables easier interpretation of the coefficients as percentage changes, since the differences factor out the constant terms, and β_3 can therefore be interpreted as the percentage change in the dependent variable (Y) as a result of a one-unit change in a predictor variable (x). The source EUI was used to incorporate all transmission, delivery, and production losses for energy used by buildings.

Overall, 3710 buildings are present in our balanced panel dataset for the years from 2011 to 2014 (inclusive). Among them, 26.9% of buildings are publicly-owned (999 buildings), and the remaining buildings are privately-owned (2711 buildings). Energy Star scores are only available for certain types of properties such as commercial buildings, so among the publicly-owned buildings, 426 properties have scores while 573 properties do not have such scores. Similarly, among the privately-owned properties, 309 buildings are scored properties, while the remaining 2402 buildings are non-scored properties (mostly multifamily buildings). From these buildings, we then identified comparable groups of office buildings, including a treatment group of 261 privately-owned scored office buildings, a first control group of 75 publicly-owned and scored office buildings, and a second control group of 39 buildings.

Table 3 provides the summary of descriptive statistics of the 375 total buildings used for the treatment and control groups in our analysis. The weather-normalized EUI ranges from 38.9 to 516.9 Kbtu/SF with a mean of 212.0 Kbtu/SF. Electricity, steam, and natural gas account for about 63%, 14%, and 12% of total energy use on average, respectively. In terms of building characteristics, a typical building is an 18-story building of 306,000 square feet. Most buildings were built between the 1910s and 1990s, and about 25.3% of buildings were built in the 1920s.

Finally, in late 2012, Hurricane Sandy disrupted energy distribution and therefore consumption, especially in Lower Manhattan. To account for the impact of Hurricane Sandy in our analysis, a binary variable was created to indicate buildings located below 34th street in Manhattan. About 9% of total observations or 36% of the buildings were affected by this storm.

Results

The estimation results from our difference-in-differences study design are displayed in Table 4 and 5. These tables indicate the overall benchmarking effect and as well as the particular effect of disclosing Energy Star scores. Significant variables that determine the building energy use are also discussed further below.

To summarize the results, this study indicates that there is a significant reduction of energy use associated with the total benchmarking policies and as well as the disclosing of Energy Star scores. However, this reduction is only statistically significant in the third and fourth year of the city's benchmarking ordinance, and each model is discussed further below.

Total benchmarking effects

Comparing the treatment and the first control group (Table 4), the coefficient of the interaction term between treatment dummy and the Year 2013 dummy is significantly negative: -0.0599. This indicates that the benchmarking policy decreased the source EUI by 6% by three years later, when controlling for building characteristics, energy use percentage by type, and effects from Hurricane Sandy. Similarly, the total effects of benchmarking policies in 2014 was a 14.3% reduction as compared to the initial year 2011. Since this effect only appears in 2013 and 2014, energy saving related to benchmarking policies appears to be only significant after three years of policy implementation. This interpretation is reasonable, as time is required for building owners and potential tenants/buyers to understand the building energy information and to incorporate it into their decision-making process.

Table 3: Descriptive Statistics Summary (N=1500 observations; or 375 buildings)

Variable name (Unit)	Mean	Std.	Min.	Max.
WNSourceEUI(kBtu/SF)	212.010	75.883	38.9	516.9
Proportion of fuel type				
Electricity use (Baseline)	0.625	0.204	0.003	1
Nature gas use	0.124	0.206	0	0.943
District steam use	0.138	0.201	0	0.997
Fuel oil no.2 use	0.019	0.093	0	0.803
Fuel oil no.4 use	0.029	0.112	0	0.809
Fuel oil no.5.6 use	0.065	0.168	0	0.935
Building features				
No. of floors	17.542	12.686	1	85
Building area(ft^2)	306,183.2	366,916.5	400	2,812,739
Built year	1939	34.062	1600	2007
Other				
Privately-owned dummy	0.800	0.400	0	1
Energy score dummy	0.896	0.305	0	1
Sandy effect dummy	0.089	0.285	0	1

Effect of disclosing Energy Star scores

Comparing the treatment and the second control group (C2) investigates the effect of disclosing particular energy information, i.e., Energy Star scores. Table 5 below shows the results. The statistically significant effect for 2013 is -0.0855 or -8.55%; this indicates that in the third year of policy implementation, the disclosure of Energy Star Scores to the public decreased the source EUI by -8.6%, when controlling for other key building features. The coefficient for the interaction term between Energy score dummy and the Year 2014 is also significantly negative, approximately -12.9% as compared with the initial year 2011. Results indicate that the reduction of source EUI associated with the Energy Star scores disclosure is significant only in 2013 and 2014, the third and the fourth year of policy implementation.

Table 4: The Effect of Disclosing Energy Score and Basic Energy Information (Treatment group T, n=261 buildings; control group C1, n=75 buildings)

Variable name	Coefficient	Sig.	Clustered Std. err.	P-value
(Intercept)	1.5702		1.0836	0.1476
Disclosure effect				
New-disclosure dummy * 2012	0.0253		0.0205	0.2166
New-disclosure dummy * 2013	-0.0599	*	0.0334	0.0729
New-disclosure dummy * 2014	-0.1426	***	0.0452	0.0016
Year dummy				
2012	-0.0132		0.0114	0.2467
2013	0.0040		0.0288	0.8904
2014	0.0440		0.0393	0.2624
Building features				
Building area (1,000,000)	0.0724		0.0560	0.1956
No. of floors	0.0028		0.0021	0.1866
Built year	0.0019	***	0.0006	<0.001
Energy use percentages				
Nature gas use	-0.0018		0.1210	0.9883
District steam use	0.0954		0.1235	0.4402
Fuel oil no.2 use	-0.5202	***	0.1379	<0.001
Fuel oil no.4 use	-0.3914	***	0.1439	0.0066
Fuel oil no.5.6 use	-0.4874	***	0.1827	0.0077
Other				
Sandy effect dummy	-0.0289		0.0347	0.4051
New-disclosure dummy	0.0081		0.0662	0.9023

Significance levels: *p<0.1; **p<0.05; ***p<0.01; the standard errors are clustered by the building ID.

One note on interpretation: it is possible that the reduction in EUI due to Energy Star scores is greater than the sum of all benchmarking policies because we derived these results using two different control groups and therefore comparisons. These two results should therefore be considered to be independent of one another.

We include key building features such as the number of floors, built year, and as well as percentages of fuel types (including use of district steam and fuel oils #2, #4, #5 and #6) in the estimation as well. As expected, other factors such as the percentages of various fuel oils are also significant in determining building energy use.

Table 5: Results of Quantifying the Effect of Disclosing Energy Score (Treatment group T, n=261, control group C2, n=39)

Variable name	Coefficient	Sig.	Clustered Std. err.	P-value
(Intercept)	1.1461		1.2342	0.3533
Disclosure effect				
Energy score dummy * 2012	-0.0647		0.0429	0.1318
Energy score dummy * 2013	-0.0855	**	0.0434	0.0493
Energy score dummy * 2014	-0.1292	***	0.0395	0.0011
Year dummy				
2012	0.0758	*	0.0401	0.059
2013	0.0203		0.0424	0.6323
2014	0.0315		0.0350	0.3674
Building features				
Building area (1,000,000)	0.0791		0.0556	0.1553
No. of floors	0.0023		0.0021	0.2809
Built year	0.0791		0.0556	0.1553
Energy use percentages				
Nature gas use	-0.2397		0.1554	0.1233
District steam use	-0.0267		0.1311	0.8387
Fuel oil no.2 use	-0.6856	***	0.1349	<0.001
Fuel oil no.4 use	-0.5955	***	0.1261	<0.001
Fuel oil no.5.6 use	-0.6148	***	0.1468	<0.001
Other				
Sandy effect dummy	-0.0132		0.0365	0.7190
Energy score dummy	0.0308		0.0705	0.6625

Significance levels: *p<0.1; **p<0.05; ***p<0.01; the standard errors are clustered by the building ID.

Limitations of this study

This study has only evaluated whether or not benchmarking policies affect energy consumption. Advocates have argued that benchmarking policies as an energy efficiency strategy could have multiple impacts, including energy saving, reduced greenhouse gas emission reductions, increased real estate valuation, and the creation of associated jobs (DOE 2015). Further examination of these other indirect disclosure effects will be important to evaluate further how benchmarking policies work. Also, this analysis is only based on four years of data. Longer-term studies will be needed to understand if benchmarking policies are effective and persistent over longer times.

The main conclusions of the paper are based on the difference-in-differences model. The buildings selected for the treatment and control groups and analyzed in this study belong to different groups because of the ways that the policy was implemented. It would be nicer to do another analysis, perhaps in another city, by implementing a randomized control trial (RCT) rather than applying a quasi-experimental design to observational data as we have done here.

Finally, this analysis was conducted only for the City of New York. However, this provides a general method for other cities to evaluate their benchmarking efforts, if they also have a policy implementation process that occurs in phases.

Conclusions and Implications

Benchmarking policies are rapidly spreading among cities, but we draw the following conclusions about this particular implementation in NYC. First, this benchmarking and disclosure policy did save energy among individual buildings in NYC by about 6% three years after implementation, and by 14% four years after the policy took effect. Similar energy policies should therefore be encouraged and promoted in more cities and regions, especially those regions in which energy consumption of buildings account for a large share of total urban energy use. Second, it takes a certain amount of time for the energy savings associated with disclosure policies to take effect: energy consumption only significantly decreased in 2013 and 2014, the third and fourth year after NYC adopted this policy. This is consistent with previous theory or conceptual studies on building energy policy (Palmer and Walls 2014; DOE 2015): market actors need some time to understand newly disclosed building energy use information and to incorporate this into their decision making processes. Efforts focusing on increasing public awareness of energy efficiency and facilitating transparency might also help to change market structure. Third, disclosure of Energy Star scores is a significant factor in the reduction of energy use, though a previous study suggests that there is still room for improvement in how scores are calculated and reported (Hsu 2014). In future studies, it would be good to understand further the processes by which market actors use and act on this information, to improve both policy designs and scoring systems.

References

Andaloro, A., R. Salomene, G. Loppolo, and L. Andaloro. 2010. "Energy certification of buildings: A comparative analysis of progress towards implementation in European countries." *Energy Policy* 38(10): 5840–5866.

Beerepoot, M. and N. Beerepoot. 2007. "Government regulation as an impetus for innovation: Evidence from energy performance regulation in the Dutch residential building sector." *Energy Policy* 35(10): 4812–4825.

Brown, R., C. Webber, and J.G. Koomey. 2002. "Status and future directions of the Energy Star program." *Energy* 27(5): 505–520.

Burr, A., C. Keicher, and D. Leipziger. 2011. *Building Energy Transparency: A Framework for Implementing U.S. Commercial Energy Rating and Disclosure Policy.* Accessed May 6, 2016. Institute for Market Transformation. http://www.buildingrating.org.

DOE (Department of Energy). 2015. *Benchmarking & Transparency Policy and Program Impact Evaluation Handbook*. Prepared for the U.S. Department of Energy by Navigant Consulting, Inc.and Steven Winter Associates, Inc.

Duflo, E., R. Glennerster, and M. Kremer. 2007. "Chapter 61 Using Randomization in Development Economics Research: A Toolkit." In T. P. S. and J. A. Strauss, ed. *Handbook of Development Economics*. Elsevier: 3895–3962.

EPA (U.S. Environmental Protection Agency). 2012. *Benchmarking and Energy Savings*. http://www.energystar.gov/ia/business/downloads/datatrends/DataTrends_Savings_20121002.pdf.

Hinge, A. and C.P. Miclea. 2015. *Empowering Energy Efficient Cities through Data*. C40 Cities Climate Leadership Group.

Hsu, D. 2014. "Improving energy benchmarking with self-reported data." *Building Research & Information* 42(5): 641–656.

Kontokosta, C.E. 2013. "Energy disclosure, market behavior, and the building data ecosystem: Market behavior and the building data ecosystem." *Annals of the New York Academy of Sciences* 1295(1): 34–43.

NRDC (Natural Resources Defense Council) and IMT (Institute for Market Transformation). 2015. *City Energy Project*. Assessed May 6, 2016. http://www.cityenergyproject.org/wp-content/uploads/2015/08/City-Energy-Project-One-Pager-Aug-2015.pdf.

Palmer, K.L. and M. Walls. 2015. "Can Benchmarking and Disclosure Laws Provide Incentives for Energy Efficiency Improvements in Buildings?" *Resources for the Future Discussion Paper* (15-09). Accessed July 9, 2015. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2564251.

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

NYC (The City of New York). 2014. *New York City Local Law 84 Benchmarking Report*. Accessed July 9, 2015. http://www.nyc.gov/html/planyc/downloads/pdf/publications/2014_nyc_ll84_benchmarking_report.pdf.

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