

# Estimating energy savings and greenhouse gas emission reductions from energy rating and disclosure policies

FÁTIMA URIARTE CÁCERES<sup>1</sup>

RECIBIDO: 04/05/2018 ACEPTADO: 26/07/2018

## ABSTRACT

The purpose of this paper is to estimate the energy savings and greenhouse gas emission reductions from energy rating and disclosure policies in New York City. During the first six year of the policy, the City saw a cumulative energy savings of 10.8% and cumulative GHG reductions of 8.5%. However, these gross changes cannot necessarily be attributed to the benchmarking and transparency policies. In order to estimate the "counterfactual" scenario – what would have happened if the policy was not implemented, we use a combination of Difference-in-Difference Estimation and Propensity Score Matching. Based on this model, we conclude that the disclosure of Energy Star Scores to the public did not have a significant effect on energy savings and GHG emissions reductions for multifamily buildings.

**Palabras-claves:** Energy; benchmarking; disclosure; causal inference.

ESTIMACIÓN DE AHORROS DE ENERGÍA Y REDUCCIONES DE EMISIONES DE GASES DE EFECTO INVERNADERO A PARTIR DE POLÍTICAS DE CALIFICACIÓN Y DIVULGACIÓN DE ENERGÍA

## RESUMEN

El propósito de este documento es estimar el ahorro de energía y las reducciones de emisiones de gases de efecto invernadero de las políticas de calificación y divulgación de energía en la ciudad de Nueva York. Durante los primeros seis años de la política, la Ciudad registró un ahorro de energía acumulada del 10.8% y reducciones acumuladas de GHG del 8.5%. Sin embargo, estos grandes cambios no necesariamente se pueden atribuir a las políticas de evaluación comparativa y transparencia. Para estimar el escenario "contrafactual", ¿qué hubiera pasado si la política no se hubiera implementado?, se utiliza el método de dobles diferencias emparejadas. Con base en este modelo, concluimos que la divulgación de Energy Star Scores al público no tuvo un efecto significativo en el ahorro de energía y las reducciones de emisiones de GEI para edificios multifamiliares.

**Keywords:** Energía; evaluación comparativa; divulgación; inferencia causal.

## 1. INTRODUCTION

Over the last decade, local legislators have begun passing energy benchmarking and disclosure laws in big cities across America. These policies, aimed at increasing buildings' energy efficiency, are part of the larger "going green" and "sustainability" movements. In 2009, for example, New York City (NYC) passed their first benchmarking and disclosure law - Local Law 84. The Appendix shows the evolution of NYC benchmarking and transparency policy.

The theory of change behind these laws is two-fold. First, benchmarking policies force building owners to actually track, analyze, and compare their energy use and costs. This ensures that every building owner actually pays attention to their energy consumption and can help them identify their largest inefficiencies. Secondly, these laws also require buildings to report their energy use for inclusion in a public database. The public database allows owners to compare their buildings' energy use to that of similar buildings and see how they stack up. Additionally, potential renters, buyers, and investors are able to view a building's energy efficiency and can use that information when assessing investment opportunities. Lawmakers theorize that by tying the building's market value to its energy efficiency, owners will be more incentivized to increase efficiency (Palmer, Walls, Gordon & Gerarden, 2013; Stavins, 2003; Winston, 2008).

The Environmental Protection Agency (EPA) has reported that these policies have decreased energy consumption by 7% over a three-year period. However, their study failed to prove actual causality between the policy and the decrease. To date, various other studies have also shown conflicting results as to the effectiveness of benchmarking policies (Meng, Hsu & Han, 2017). Nevertheless, in 2016, New York City expanded their benchmarking and disclosure requirement to include buildings over 25,000 feet with Local Law 133. Additionally, in 2017, New York City approved a bill requiring large buildings to post their energy efficiency scores, and associated grades, near their public entrances. This is intended to make information about a building's energy use easily available to tenants and visitors, including prospective purchasers or lessees, and presented in the easily understood form of a grade.

<sup>1</sup> Master of Public Administration and Master of Mathematical Statistics, Independent Researcher. E-mail: [fatima.uriarte@columbia.edu](mailto:fatima.uriarte@columbia.edu)

Before the city continues to use resources to expand this policy further, it is critical to perform an evaluation of it (Delmas Montes-Sancho & Shimshack, 2010; Hsu, 2014; Granade, Creyts, Derkach, Farese, Nyquist & Ostrowski, 2009). Since new data has been reported as a result of the 2016 law, it is an opportune time to analyze if energy benchmarking and disclosure laws actually cause a decrease in energy consumption and GHG emissions. Proof of the policy's utility will be beneficial to advocate for greater expansion in New York City as well as push for reform in other cities considering these policies. On the other hand, evidence that these policies are ineffective can be used as cautionary tales for other cities before they waste resources pushing these policies.

Furthermore, since the New York City benchmarking and disclosure policy has multiple elements (disclosure, benchmarking, scoring), it is important to assess which aspects are the main causal change mechanisms. By understanding which particular features of the policy are most effective, New York City can better tailor it and more effectively allocate resources to promote the most significant change. Therefore, this research will investigate the overall gross impacts of the benchmarking and disclosure policy as well as the causal effectiveness of one aspect of the policy - Energy Star Scores - on changes in energy consumption and GHG emissions.

## 2. BACKGROUND AND CONTEXT

Figure 1 provides a description of how the NYC Benchmarking and Disclosure Policy is supposed to deliver the desired results. In general, immediate and short-term outcomes focus on the earliest effects of awareness and the initial actions taken towards saving energy. Intermediate outcomes are the intended changes to market structure or market actor behavior, such as owners making building improvements to secure deeper energy savings and property values incorporating energy performance. Long-term outcomes are the intended market effects expressed in the persistent energy savings and greenhouse gas reductions (US Department of Energy, 2015).

In order to achieve these objectives, many forms of information disclosure are taken place, such as:

Reports: The New York Mayor's Office of Sustainability releases on its website annual reports, explaining in details the updated program and its results.

Databases in NYC Open Data: In the public database NYC Open Data, there are data of NYC municipal and private buildings for the calendar years 2011-2016.

Energy Star Score: This is a 1-to-100 percentile ranking, calculated in Portfolio Manager, based on

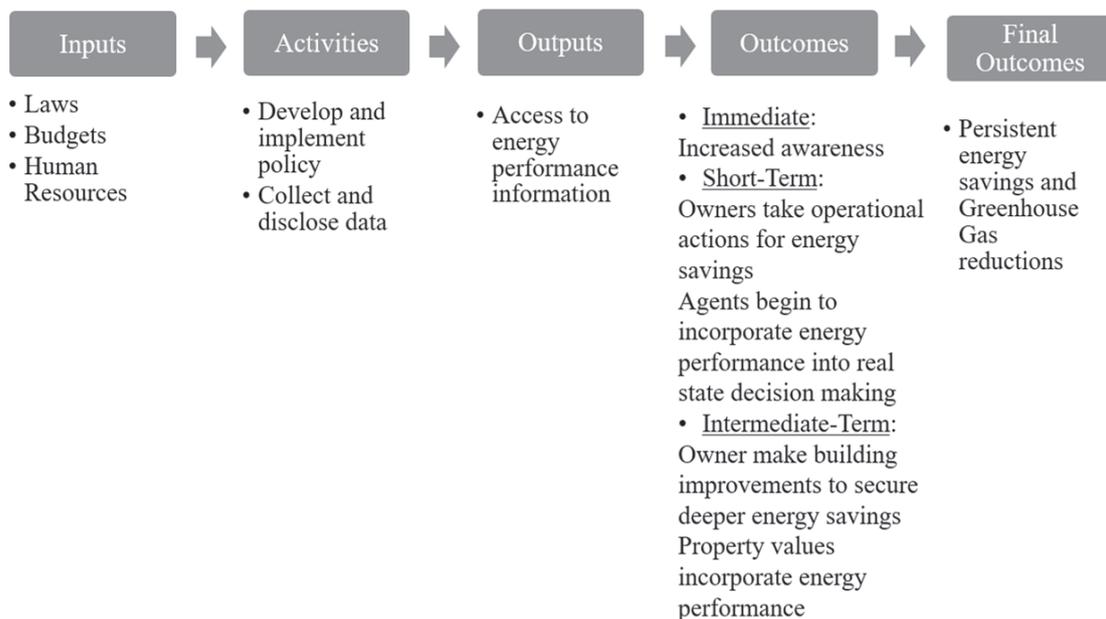


Figure 1. Theory of Change

Source: Prepared by the author, based on Benchmarking Reports (Mayor's Office of Long-Term Planning & Sustainability, 2012, 2013, 2014; Urban Green Council, 2016, 2017), and Benchmarking and Transparency Policy Report – Impact Evaluation Report (U.S. Department of Energy, 2015).

self-reported energy usage for the reporting year. A higher score indicates better performance.

User friendly websites: Some websites or platforms are available for the public to check the data using visualization tools. For example: “Energy Star Portfolio Manager”, and “Metered New York”.

### 3. METHODOLOGY

#### 3.1. Data

Our primary sources are the databases available in NYC Open Data that cover the period 2011 to 2016. We use weather normalized source energy use intensity (EUI) and total GHG emissions as dependent variables. The Mayor’s Office of Long-Term Planning & Sustainability (2017) defines the weather normalized source EUI as the “energy use intensity as calculated by the source of energy generation in kBtus per gross square foot (kBtu/ft<sup>2</sup>) for the reporting year, normalized for weather” (p.3), and GHG emissions as “the total direct and indirect greenhouse gases emitted by the property, reported in metric tons of carbon dioxide equivalent (Mt-CO<sub>2</sub>e) for the reporting year” (p.3).

Additionally, we use Energy Star Scores if they apply, and other building characteristics available in the City of New York’s Primary Land Use Tax Lot Output (PLUTO) dataset.

Before estimating the energy impact of NYC benchmarking and transparency policies, we do a series of data cleaning steps following the methodology developed by the US Department of Energy (2015), such as: (i) removed all duplicate entries, (ii) removed all buildings that were not available in PLUTO dataset, (iii) removed all building with building types that were zero or blank, (iv) removed all buildings with site EUIs outside of a reasonable range (5-1,000 kBtu/ft<sup>2</sup>), and (v) removed all buildings with abnormal changes in site EUIs between two years being compared <sup>2</sup>.

#### 3.2. Research Design

According to the US Department of Energy (2015), benchmarking and transparency policies have the following impacts:

- Gross Impacts: The change in buildings’ energy usage over time inclusive of actions taken to reduce energy consumption, as well as their participation in other energy efficiency activities or programs.
- Net Impacts: The subset of measured gross energy changes attributable to the Benchmarking and Transparency policy. That is, the net savings after taking into account natural market forces and the impacts from other local, state, federal, and utility energy-efficiency program and tax credit initiatives.

The gross impacts are calculated by comparing the same population of buildings over the years. However, this evaluation cannot necessarily attribute these energy savings to the benchmarking and transparency policies. In theory, the actual energy savings achieved by a benchmarking and transparency policy should be equal to the difference between the amount of energy used by buildings subject to the policy and the amount of energy they would have used had the policy not been adopted. This baseline is called the “counterfactual” scenario – what would have happened if the policy was not implemented. As the counterfactual scenario does not exist and cannot be directly measured, defining an approach that approximates it is the fundamental challenge to estimating energy impacts and documenting the benefits of policies (US Department of Energy, 2015). As such, we are only able to isolate the effects of Energy Star Scores on consumption behavior. Net impacts for Energy Star Scores will be calculated using a combination of propensity score matching and difference and difference techniques.

#### Propensity Score Matching

The standard theoretical framework for analyzing the problem of impact evaluation is based on the potential outcome model or Roy-Rubin model (Bernal & Peña, 2012). Formally, we define the treatment indicator as  $D_i$ , with  $D_i = 1$  if the individual  $i$  receives the treatment and zero otherwise. The outcome variables are defined as  $Y_i(D_i)$  for each individual  $i = 1 \dots N$  and  $N$  denotes the total population. That is,  $Y_i(1)$  is the outcome variable if the individual  $i$  is treated and  $Y_i(0)$  is the outcome variable if the individual  $i$  is not treated. The effect of the treatment, or impact of an event on a variable of interest, for an individual  $i$  can be written as:

$$\tau_i = Y_i(1) - Y_i(0)$$

The fundamental problem of impact evaluation is that in reality only one of the two potential results

<sup>2</sup> Specifically, buildings that had site EUIs that increased or decreased more than 50% were removed because changes beyond these thresholds were likely due to reasons other than the policy (change in occupancy, erroneous data entry, change in space usage, etc.)

$Y_i(1)$  or  $Y_i(0)$  is given for each individual  $i$ , but never both, so the expected result can be written as:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0) = \{ Y_i(1) \text{ if } D_i = 1 \text{ or } Y_i(0) \text{ if } D_i = 0$$

Because one of the results in the above equation is not observable for the same individual, it is not possible to estimate the individual effect of the treatment  $T_i$ . Therefore, the analysis should focus on the average impact of treatment on the population or subsets of the population. In the present study, we want to estimate the average treatment effects on the treated (ATT); that is, the average effect of NYC benchmarking and transparency policies on energy consumptions and GHG emissions of the subset of buildings that actually disclosure its Energy Star Score, which corresponds to the difference between the average of the energy consumption (and GHG emissions) in the group of buildings subject to Energy Star policy and the average of the energy consumption (and GHG emissions) that these buildings would have used if the policy was not implemented.

$$\tau_{ATT} = E(\tau_i / D_i = 1) = E[Y_i(1) / D_i = 1] - E[Y_i(0) / D_i = 1]$$

Where  $E[\cdot / D]$  denotes the operator of conditional expectations.

In this case  $E[Y_i(1) / D_i = 1]$  is the expected value of the outcome variable in the treatment group in the presence of the treatment and  $E[Y_i(0) / D_i = 1]$ , which is known as counterfactual, is the expected value of the outcome variable in the treatment group in the absence of treatment. Obviously, this is a hypothetical result, so it is not observed and therefore is not recorded in the data.

One way of estimating the counterfactual is to assume that, given a set of observable variables  $X$  that are not determined by the treatment, the potential results,  $Y_i(0)$  and  $Y_i(1)$ , are independent of the treatment allocation. This assumption implies that all the variables that simultaneously affect the allocation to the treatment and the potential results ( $Y_i$ ) are observed by the researcher and included in the model that is estimated. This estimation method is known in the impact assessment literature as matching.

Suppose we want to estimate the impact of the intervention by comparing the outcome variables of the treatment and control group:

$$E[Y_i(1) / D_i = 1] - E[Y_i(0) / D_i = 0]$$

Adding and subtracting the counterfactual,  $E[Y_i(0) / D_i = 1]$ , we have:

$$= E[Y_i(1) / D_i = 1] - E[Y_i(0) / D_i = 1] + E[Y_i(0) / D_i = 1] - E[Y_i(0) / D_i = 0]$$

The first two terms are the average treatment effects on the treated ( $T_{ATT}$ ). The last two measure the selection bias: the difference between the counterfactual and the outcome variable for the control group.

One possibility is to assume that the selection bias is due only to differences in the observable characteristics, so that:

$$Y(0), Y(1) \perp D / X, \quad \forall X$$

What is called the conditional independence (CI) assumption. This assumption ensures that by conditioning in the observable variables  $X$ ,  $E[Y_i(0) / D_i = 1, X] = E[Y_i(0) / D_i = 0, X]$  and, therefore, that the selection bias is equal to zero. Thus, calculating the impact of the event of interest as the difference in the average of the outcome variables of the treatment and control group, conditioned on the observed variables, generates an unbiased estimate of the true effect of the treatment.

Note that the implementation of the estimator can be computationally complex if there is a large number of variables based on which we are matching, that is, if the vector  $X$  has a very large dimension. One possibility is to match individuals based on their estimated propensity score, given their observable characteristics  $P(X)$ , where:

$$P(X) = P(D=1|X)$$

That is, each individual of the treatment group will be matched with that of the control group with a propensity score in the program sufficiently close. All pairing estimators contrast the result variable of a treated individual with the results of one or more members of the control group that most closely resemble the individual treated, based on the measure  $P(X)$ .

The common support condition (CS) implies that only individuals from the control group that have propensity scores  $P(X)$  similar to the propensity scores of the treatment group are used in the estimation, which is known as propensity score matching (PSM). For example, if there are individuals from the control group with very low propensity score, but no treated individual exhibits such a low propensity score, then these individuals from the control group will be discarded when making the pairing.

Assuming that the conditions of CI and CS are met, the ATT estimator by PSM would be given by:

$$\tau_{ATT}^{PSM} = E_{P(X)|D=1}\{E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)]\}$$

Where  $E_{P(X)|D=1}$  is the expected value with respect to the propensity score  $P(X)$ , conditional on being a participant of the program.

The result of the pairing method is reliable as long as there are reasons to believe that variables not observable or not available in the database are not a fundamental determinant of potential results.

**Difference-in-Difference Matching**

One of the advantages of the matching method is that it can be applied to a single information survey, as long as there are observations of treatments and controls. However, if there is longitudinal information, or repeated cross sections, the matching method can be combined with the difference-in-differences (DiD) method, in such a way that the assumptions of each of these methods are relaxed, and therefore, the results become more robust. Thus, through the combination of matching and DiD we have eliminated the time constant unobserved effects.

By having longitudinal information and combining the PSM with differences in differences, the selection to treatment may also depend on unobservable variables, as long as they remain constant over time. If information is available for two periods,  $t=1$  is the period before the implementation of the treatment and  $t=2$  is the period after the implementation of the treatment or follow-up period. The condition of conditional independence in this context is re-defined as the assumption that the evolution of the non-observables is independent of the treatment, when controlling for:

$$u_{i2} - u_{i1} \perp D/X, \forall X$$

The subscript of  $u$  indicates the period to which that observation corresponds. For example,  $u_{i1}$  corresponds to the observation  $u$  in the period prior to the implementation of the treatment. This version of conditional independence implies that the treatments and controls would have evolved in the same way if neither of them had been the object of the treatment. That is, the objective variable for both groups follows a common trend.

Since we are combining the matching method with the difference-in-difference method, we need to restrict common support to ensure that the treatment

and control groups are very similar and thus identify the ATT appropriately.

Assuming that the conditions of CI and CS are met, the ATT estimator of difference-in-difference matching (DDM), in the common support region, would be given by:

$$\tau_{ATT}^{DDM} = E_{P(X)|D=1}\{E[Y_2(1) - Y_1(1)|D = 1, P(X)] - E[Y_2(0) - Y_1(0)|D = 0, P(X)]\}$$

Where  $E_{P(X)|D=1}$  is the expected value with respect to the propensity score, conditional on being a participant of the program. The subscript of  $Y$  indicates the period to which that observation of the result variable corresponds. For example,  $Y_1$  corresponds to the observation of  $Y$  in the period before treatment. Finally, the condition  $.|D$  indicates whether the observation corresponds to an individual in the treatment group ( $D = 1$ ) or to an individual in the control group ( $D = 0$ ).

**4. RESULTS**

**4.1. Gross Impacts**

First, gross energy impacts were calculated by comparing the same set of 582 buildings that reported their total EUI every year from 2011 to 2016. Overall energy savings were calculated by comparing the difference in total EUIs (Energy Use Intensity/Square foot) between 2011 and each subsequent year. Results indicate that total EUIs were reduced each year with a cumulative reduction of 10.8% from 2011 to 2016 (Figure 2). This equates to an average savings of 3,600 MMBtus per building.

However, Figure 3 shows that these results varied significantly when evaluating by building type. Multifamily housing, for example, saw an increase in consumption of 4.82% over the years while Offices and Hotels saw a decrease.

Similar to energy savings, gross GHG impacts were calculated by taking the difference in the total GHG emissions between each year from the same group of 318 buildings that reported annually. While we see a cumulative decline of 8.5% in GHG emissions from 2011 to 2016, there was not a monotonic decrease throughout the years (Figure 4). There was a 20.9% drop between 2011 and 2012 but then total GHG emissions fluctuate between increasing and decreasing until 2016. Looking at just 2011 and 2016, there was a difference of 15,806 MtCO2e which equates to a reduction of 50 MtCO2e per building. It is worth mentioning that there was not sufficient data to analyze by building type.

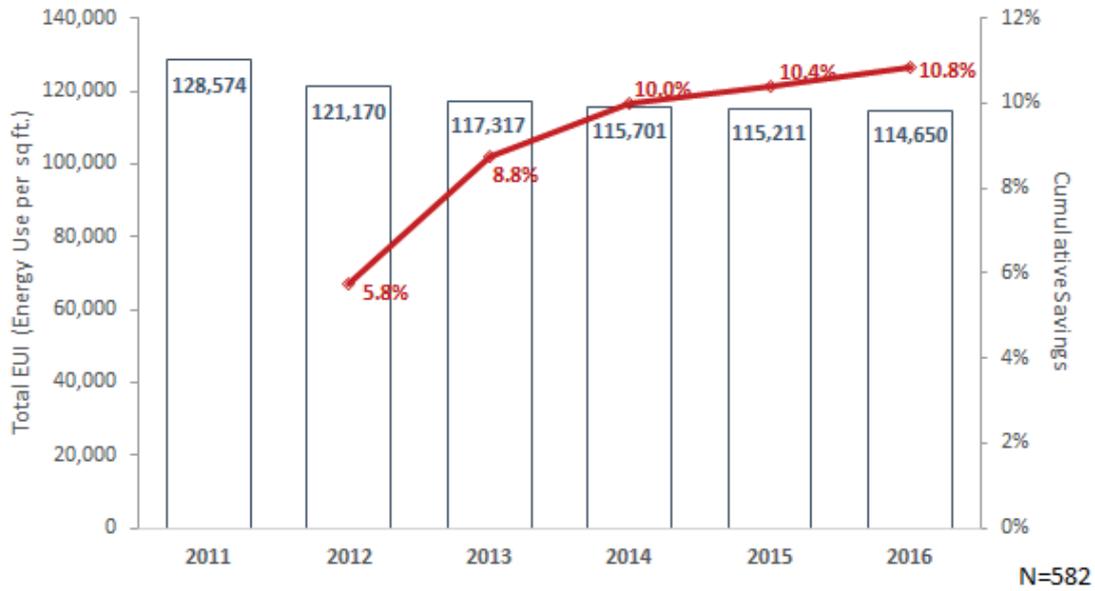


Figure 2. Gross Energy Savings

Source: Prepared by the author, based on the Energy and Water Data Disclosure Dataset.

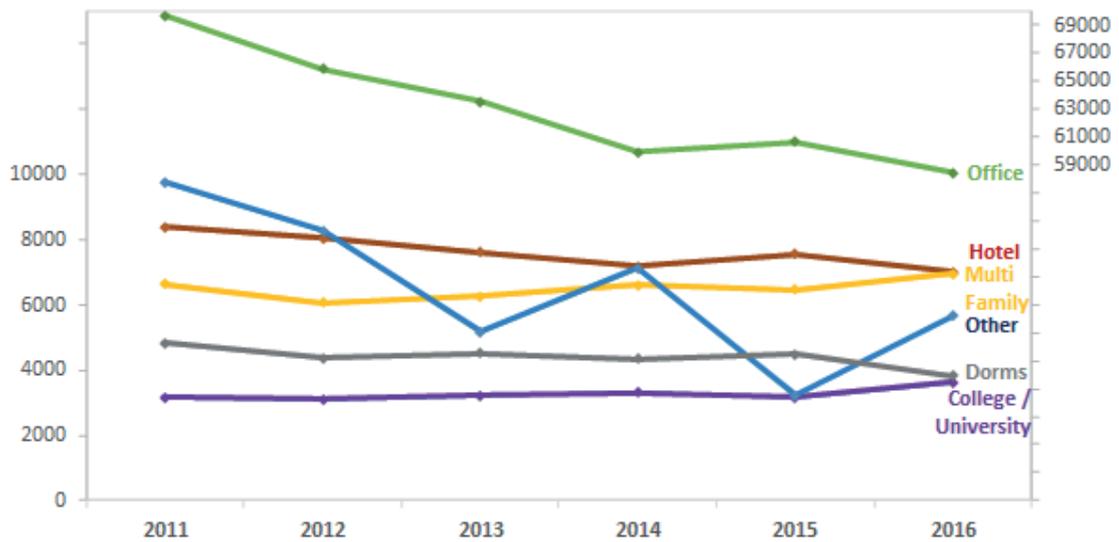


Figure 3. Energy Consumption by Building Type

Source: Prepared by the author, based on the Energy and Water Data Disclosure Dataset.

While we see reductions in both Energy use and GHG emissions over the 6-year time-period, net impacts need to be calculated to determine if these reductions can be attributed to the Benchmarking and Disclosure policy.

#### 4.2. Net Impacts: Difference-in-Difference Matching

This section investigates the effect of disclosing the Energy Star Score. It is worth mentioning that the variable to be estimated is the difference in the

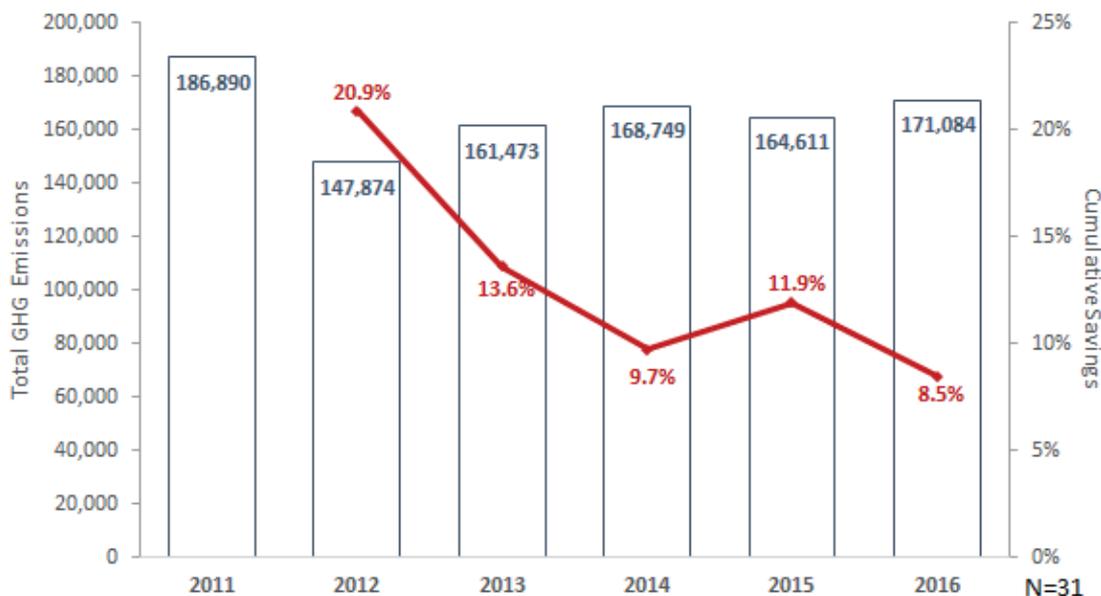


Figure 4. Gross GHG Emissions Savings

Source: Prepared by the author, based on the Energy and Water Data Disclosure Dataset.

Table 1. Effects of Disclosing Energy Star Scores (percentage points)

	Treated	Controls	Difference	S.E.	T-stat
WNSourceE UI	3.26	3.83	-0.58	1.93	-0.3
GHG Emissions	1.86	1.52	0.35	2.61	0.13

Source: Prepared by the author, based on the Energy and Water Data Disclosure Dataset and the City of New York's Primary Land Use Tax Lot Output (PLUTO).

weather normalized source energy use intensity and GHG emissions, before and after the Energy Star disclosure. Thus, each difference in those variables in the treatment group will be paired with a set of controls, according to the nearest neighbor matching<sup>3</sup>.

In this exercise we use multifamily housing buildings that have information before and after the implementation of the policy. We chose to use only multifamily buildings over other types of buildings such as offices or hotels because there may be factors that make energy consumption in the residential sector different from energy consumption in the commercial or industrial sector. Additionally, multifamily buildings represent 74% of total buildings in NYC in 2016.

To estimate the effect of Energy Star Scores using DiD matching, we first estimate the propensity score. Since this variable is binary, it can be estimated using logit or probit models. Most of the variables included are highly significant and the model is significant as a whole. With the estimated coefficients, the propensity score is predicted for both the treatment and control groups. It is worth mentioning that the common support region in this application is approximate 5,000 buildings<sup>4</sup>.

Table 1 shows the impact of disclosing Energy Star Scores on the variables of interest estimated by DiD matching. The disclosure of Energy Star Scores to the public decreased the source EU by 0.58 percentage points. Additionally, it seems to be an increased in GHG emissions of 0.35 percentage

3 The PSM estimator for nearest neighbor consists of taking each individual from the treatment group and looking for an individual in the control group with the closest propensity score, that is, the nearest neighbor based on P(X). That is, the individual in the control group such that the distance between their predicted propensity score of the individual in the treatment group is minimal.

4 In order to test the robustness of the method, the space of the estimated probabilities of the treatment and control groups was divided into blocks of propensity scores. The strata were chosen with the objective of balancing the propensity scores among them, that is, the average propensity score of the treatment and control groups was equal. In this way, an optimal number of ten blocks was determined and it was proved, block by block, that there were no statistically significant differences between buildings of the treatment and control group.

points. However, these impacts are not statistically significant. From the results obtained, it seems that the Energy Star Scores have not attracted the attention of market actors to energy use.

### **Limitations**

The main conclusion of this paper relies on the DiD Matching model. However, as was stated earlier in the document, we assume that selection to treatment not only depends on observable variables, but also on unobservable variables, as long as they remain constant over time. There may be the case that some unobservable variable change over time, affecting our results. Also, this analysis is based only on three years of data because Energy Star Scores for multifamily buildings were available since 2014, so our results may indicate that actors and the market require some time to incorporate such energy ranking information.

Our result is different from the one obtained by Meng et al. (2017). After applying a novel difference-in-difference strategy, they conclude that in the fourth year of policy implementation, the disclosure of Energy Star Scores to the public decreased the source EUI by 12.9%, when controlling for other key building features. However, they analyzed the energy consumption of private office buildings using public buildings as controls, while we analyzed multifamily buildings.

### **5. CONCLUSION**

The gross energy savings increased steadily between 2010 and 2016 while reductions in GHG emissions fluctuated over the years but saw an overall decrease when comparing 2011 to 2016. However, the actual energy savings achieved by a benchmarking and transparency policy should be equal to the difference between the amount of energy used by buildings subject to the policy and the amount of energy they would have used had the policy not been adopted. In order to estimate the “counterfactual” scenario – what would have happened if the policy was not implemented, we use a combination of Difference-in-Difference Estimation and Propensity Score Matching.

Our quasi-experimental evaluation determined that we cannot necessarily attribute these energy savings to the benchmarking and transparency policies. That is, the net savings after taking into account natural market forces and the impacts from other local, state, federal, and utility energy-efficiency program and tax credit initiatives is not statisti-

cally significant. Therefore, we conclude that the disclosure of Energy Star Scores to the public did not have a significant effect on energy savings and GHG emissions reductions for multifamily buildings in the first years of policy implementation.

It seems that the Energy Star Scores have not attracted the attention of market actors to energy use. Although Local Law 84 requires that property owners report energy performance data to the City, which then publishes the data online, it does not require property owners to ensure that tenants are at any point presented with the information collected. We suggest to make the information about energy use easily available to tenants. There is evidence that benchmarking and transparency policies work in Latin America, but these countries have focus their policies in providing information to the individuals that are the ones who actually consume the service (Scartascini, 2017). The information is provided in their bills using behavioral economic insights<sup>5</sup>.

### **6. REFERENCES BIBLIOGRAPHIC**

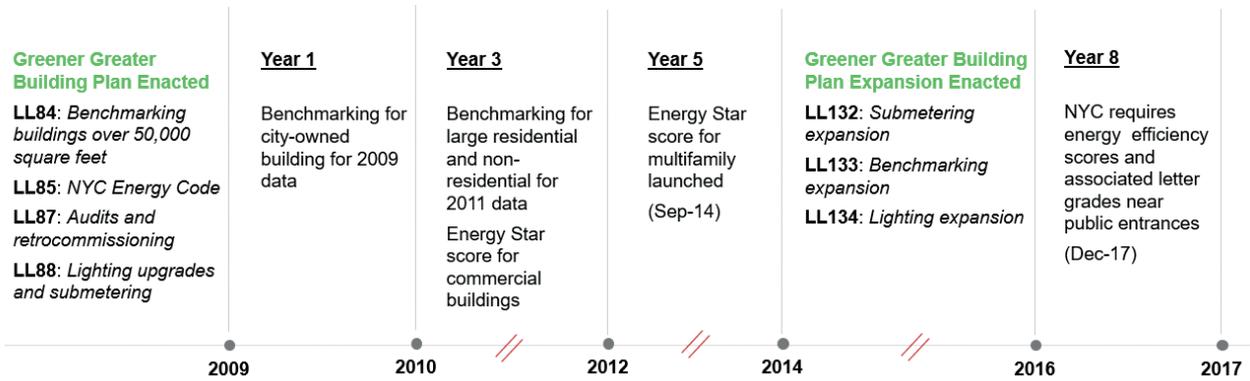
- [1] Bernal, R., & Peña, X. (2011). *Guía práctica para la evaluación de impacto*. Ediciones Uniandes-Universidad de los Andes.
- [2] Datta, S., Miranda, J. J., Zoratto, L., Calvo-González, O., Darling, M., & Lorenzana, K. (2015). *A behavioral approach to water conservation: evidence from Costa Rica*. The World Bank.
- [3] Delmas, M., Montes-Sancho, M. J., & Shimshack, J. P. (2010). Information disclosure policies: Evidence from the electricity industry. *Economic Inquiry*, 48(2), 483-498.
- [4] Granade, H. C., Creyts, J., Derkach, A., Farese, P., Nyquist, S., & Ostrowski, K. (2009). *Unlocking energy efficiency in the US economy*. McKinsey & Company.
- [5] Hsu, D. (2014). How much information disclosure of building energy performance is necessary?. *Energy Policy*, 64, 263-272.

5 For instance, Datta, Miranda, Zoratto y Calvo-Gonzalez (2015) present the design a set of three simple and replicable behavioral interventions, which use stickers that can be added to water bills at low cost, and test their impact on water consumption in Belen, Costa Rica, using a randomized control trial. Two of the three interventions were found to decrease water consumption significantly in the months following the intervention. They found that raising awareness about how much water an individual consumes, and comparing this consumption level with peers, can go a long way in helping change individuals' behavior regarding the use of a finite resource such as water.

- [6] Mayor's Office of Long-Term Planning & Sustainability. (2012). *New York city local Law 84 benchmarking report*.
- [7] Mayor's Office of Long-Term Planning & Sustainability. (2013). *New York city local Law 84 benchmarking report*.
- [8] Mayor's Office of Long-Term Planning & Sustainability. (2014). *New York city local Law 84 benchmarking report*.
- [9] Mayor's Office of Long-Term Planning & Sustainability. (2017). *Benchmarking Data Disclosure Definitions*.
- [10] Meng, T., Hsu, D., & Han, A. (2017). Estimating energy savings from benchmarking policies in New York City. *Energy*, 133, 415-423.
- [11] Palmer, K., Walls, M., Gordon, H., & Gerarden, T. (2013). Assessing the energy-efficiency information gap: results from a survey of home energy auditors. *Energy Efficiency*, 6(2), 271-292.
- [12] Stavins, R. N. (2003). Experience with market-based environmental policy instruments. In *Handbook of environmental economics* (Vol. 1, pp. 355-435). Elsevier.
- [13] Scartascini, C. (2017). Behavioral Economics and Innovation in Cities: Increasing Fiscal Revenues and Improving Public Policies in LAC. Inter-American Development Bank.
- [14] Urban Green Council. (2016). *New York City's energy and water use report*.
- [15] Urban Green Council. (2017). *New York City's energy and water use report*.
- [16] U.S. Department of Energy. (2015). *New York City Benchmarking and Transparency Policy Impact Evaluation Report*.
- [17] 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), 704-17.

APPENDIX

Evolution of NYC Benchmarking and Transparency Policy



Source: Prepared by the author, based on Benchmarking Reports (Mayor's Office of Long-Term Planning & Sustainability, 2012, 2013, 2014; Urban Green Council, 2016, 2017)